

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 38.5 Seconds

(without alignments)  
4564.770 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLTGALGLAFAFASWAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mbc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	425	14.9	576	Q8J2Q5	Q8J2Q5 gibberella
2	419	14.7	593	Q9C6S0	Q9C6S0 arabidopsis
3	419	14.7	595	Q93VK5	Q93VK5 arabidopsis
4	417	14.6	503	Q97689	Q97689 sus scrofa
5	417	14.6	517	Q8LIR5	Q8LIR5 oryza sativ
6	403.5	14.1	503	Q95LRS	Q95LRS sus scrofa
7	403	14.1	515	Q9LUD2	Q9LUD2 arabidopsis
8	401	14.1	515	Q8WIE1	Q8WIE1 arabidopsis
9	400.5	14.0	519	Q9ASR3	Q9ASR3 arabidopsis
10	397.5	13.9	519	Q8LHV0	Q8LHV0 oryza sativ
11	394	13.8	503	Q86SK3	Q86SK3 homo sapien
12	393.5	13.8	544	Q9FBE1	Q9FBE1 oryza sativ
13	392	13.7	497	Q8CJF2	Q8CJF2 rattus norv
14	389	13.6	492	Q8IS07	Q8IS07 mamestra br
15	387	13.6	523	Q8LQ77	Q8LQ77 oryza sativ
16	386.5	13.6	511	Q8HYV6	Q8HYV6 capra hircu

17	385.5	13.5	535	4	Q9H241	Q9H241 homo sapien
18	384.5	13.5	531	10	Q9FED3	Q9FED3 oryza sativ
19	382.5	13.4	527	10	Q8LGM8	Q8LGM8 zea mays su
20	382	13.4	430	4	Q7Z448	Q7Z448 homo sapien
21	381	13.4	503	6	Q8HZK1	Q8HZK1 canis famil
22	381	13.4	520	10	Q8LQ44	Q8LQ44 oryza sativ
23	380.5	13.3	505	10	Q9LUD0	Q9LUD0 arabidopsis
24	374	13.1	533	5	Q964T2	Q964T2 blattella g
25	371	13.0	507	10	Q8LQ38	Q8LQ38 oryza sativ
26	368.5	12.9	523	10	Q9FD21	Q9FD21 oryza sativ
27	365	12.8	512	10	Q9LUC8	Q9LUC8 arabidopsis
28	365	12.8	520	10	Q48786	Q48786 arabidopsis
29	364.5	12.8	503	5	Q8WQ89	Q8WQ89 anopheles g
30	363.5	12.7	511	10	Q64F31	Q64F31 arabidopsis
31	363.5	12.7	512	10	Q9LUD3	Q9LUD3 arabidopsis
32	363	12.7	502	11	Q06884	Q06884 rattus norv
33	362.5	12.7	523	10	Q9CAD6	Q9CAD6 arabidopsis
34	362	12.7	506	10	Q9LUD1	Q9LUD1 arabidopsis
35	361	12.7	528	10	Q9FBE0	Q9FBE0 oryza sativ
36	360	12.6	512	10	Q8W2N3	Q8W2N3 vicia sativ
37	359.5	12.6	512	10	Q9LUC9	Q9LUC9 arabidopsis
38	357.5	12.5	512	10	Q9LUC5	Q9LUC5 arabidopsis
39	357.5	12.5	528	10	Q8LL74	Q8LL74 zea mays su
40	356.5	12.5	523	10	Q8LAQ9	Q8LAQ9 arabidopsis
41	356	12.5	504	11	Q9EQW4	Q9EQW4 mus musculu
42	355.5	12.5	512	10	Q9ZW95	Q9ZW95 arabidopsis
43	354.5	12.4	506	10	Q7XTY4	Q7XTY4 oryza sativ
44	354.5	12.4	560	10	Q8W372	Q8W372 oryza sativ
45	354.5	12.4	560	10	Q7XD91	Q7XD91 oryza sativ

## ALIGNMENTS

RESULT 1  
Q8J2Q5 PRELIMINARY; PRT; 576 AA.  
ID Q8J2Q5  
AC Q8J2Q5  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Fum15p.  
GN FUM15.  
OS Gibberella moniliformis.  
OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;  
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
CX NCBI\_TaxID=117187;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M-3125;  
RX MEDLINE=99343882; PubMed=10413619;  
RA Proctor R.H., Desjardins A.E., Plattner R.D., Hohn T.M.;  
RT "A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in Gibberella fujikuroi mating population A.";  
RL Fungal Genet. Biol. 27:100-112(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M-3125;  
RX MEDLINE=21585628; PubMed=11728154;  
RA Seo J.A., Proctor R.H., Plattner R.D.;  
RT "Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in Fusarium verticillioides.";  
RL Fungal Genet. Biol. 34:155-165(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M-3125;  
RX Proctor R.H., Brown D.W., Plattner R.D., Desjardins A.E.;  
RT "Co-expression of fifteen contiguous genes delineates a fumonisin biosynthetic gene cluster in Gibberella moniliformis.";  
RL Fungal Genet. Biol. 0:0-0(2003).  
DR EMBL; AF155773; AAN74818.1; -;  
DR GO; GO:0003676; F:nicotinic acid binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.

```
DR InterPro: IPR001128; Cytochrome P450.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00067; P450; 1.
DR PRINTS; P00385; P450.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 576 AA; 64918 MW; B1693773CA36A8FB CRC64;

Query Match
  14.9%; Score 425; DB 3; Length 576;
Best Local Similarity 24.9%; Pred. No. 1.4e-22;
Matches 147; Conservative 100; Mismatches 230; Indels 114; Gaps 21;

QY 13 LAAFSWASIAFFSLYLAAPRSSLYNQGFN--HTNYFTGNFLDLSARTGBEHAKEYK 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 50 ISHFVWSAI-IWPTFSPRLR-----QLPNVPSDGLSKETRLVSEPRGVQSDWINS 102
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 GS-----TLRFAGIAGAPVLNSTDPKFNHVM-KEAYDPKGMARVLRIATGCVGTAE 125
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 103 SNRPVDLARYRFLGPERLLIISPKALAEVLTKSPKGLIVSELKQATGNGVLLAE 162
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 126 GSAHKRRHRIMIPISAAQVKSNVPFLEKGMELVDKMDAAEKDMAGVSGAGEKATR 185
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 GSEHKQRKALCTAFNYRIKXLYPVFNDVAGEPATVL-----EKQPTG-----TPRTS 212
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 186 LETEGVDVKGWGRATLDNALAGFDYKSDSLQNTNELYVAFVGLTDGFPAPTLDS---- 241
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 213 DTTAVDIVDWSRATLDIIGRAGMGQGFDAIQNDDSLHQAYRMI---PEPSRGAIFLA 269
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 -----EKAIWMDVPVFRYMKRHEIPLTQGLAVSRVVGIELMEQKQAVLGSASDAQVD 296
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 270 LLRLIFERLVNVLPLRRNKRHH-----GQIVRSKCQLIRERKEKI--KQKAGVD 321
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 KDVQGRDILSLVRANIANANIPESQKLSDEVLVAQISNLLFAGYETSSVLTWVHRLS 356
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 322 N---SGNDILTALLNGV-----FTDEQILDQMTFLAAGHETATALTWAIYILC 369
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 EDKAVQDKLREBEICQ-----IDTMDPTLDLN-----ALPYLEAFVKESLDDPPSP 403
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 370 KQEVQNRRLREIRHFNHPNPKWPSRPSNTLQCADFKPLYNVVCLEWRYFAPIP 429
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 404 YANRECKDEDFIAPVIGRGVINEVRITKGTWMLPLFNINRSKFTYIGSDAEFR 463
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 LTMREATCD-----TTILHTFVPAGTRIILAPRVNTRDSALWGPDANNFN 474
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 PERWLEDVTDLSNIEAPYGHQ-----ASFISGPRACGKRFVAVEMKAF 508
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 475 PDR-----NOKSDSTAAPVEVGRTEARSNVADTLFLHPSRCSIGQSFARVEFAIL 529
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 509 LFTLRRVQFEPIISHPEYEHITLIISR---PRIVGREKEGYQMLQKVPV 556
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 530 LATLIANFEFQ-IEDESLLDERNISIRGATSRIVG-----GLKVRVRPI 573
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
Q9C6S0 PRELIMINARY; PRT; 593 AA.
AC Q9C6S0;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Cytochrome P450, putative.
GN F5M6.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
```

```
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y., L.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana."
RL Nature 408:816-820(2000).
CC !- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
EMBL; AC079041; AAG50718.1; -.
DR PIR; F86441; F86441.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; P00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
KW SEQUENCE 593 AA; 66643 MW; E80CBE9B8B2BD199 CRC64;

Query Match
  14.7%; Score 419; DB 10; Length 593;
Best Local Similarity 26.4%; Pred. No. 4.1e-22;
Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NPLDILSARTG--BEHAKYREKYGST-----LRPAGI-----AGAPVLN 86
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 93 NVLDFMFDWTGSDQDPKVPKAGSIQAVRNEAFFIPLYELFTYGGIFELTFGPKSFII 152
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 STDPKVFNVHMK-AYDPKFGMAARVLRIATGCVVTAEGEAHKHRRIMPSLSAQAV 145
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 153 VSDPSIAKHILKNAKAYSK-GILAEILDFVMGKGLIPADGEIWRRRRAIVPALHQYV 211
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 KSMVPFLEKGMELVDKMDAAEKDMAGVSGAGEKKAETEGVDVKGWGRATLDVM 205
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 AAMISLFGASDRLCQKL--DAA-----ALKGEVEVMESLSRLTDII 253
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 206 ALAGFDYKSDSLQNTNELYVAFVGLTDGFPAPTLDSFKAIMWDFVYVETMKRHEIPLT 265
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 GKAVFNDFDLSLTDNTGVIEAVYTVLREADRSVPIP--VWD-IFWKDIS-----PRQ 305
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 QGLAVSRVRVGIEMEQKQKQAVLGSASDAQVDKDVQ-----GRD--ILSLVRANIAA 316
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 306 RAVATSLKLINDTLD-----LIATCKRWEBEELQFHEEYMNRPDSILHFL----- 353
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 NLPESQKLSDEVLVAQISNLLFAGYETSSVLTWVHRLSDEKAVQDKLREBEICQIDTD- 375
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 -LAGDDVSSKQLRDDLTMTLIAGHTSAALVTWTFYLLTFTEPSVAKQGEVDVIGDR 412
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 376 MPTDELNALPYLEAFVKESLRLLDPSPFYANRRLCKDDEDFIAPVIGRGVINEVR 435
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 PFTIQDMKKLKYTRVNMESLRLYPQPPVILRSI-----DNDILGEYPI 457
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 436 TKGTMVMLPLFNINRSKFTYIGSDAEFRPERWLEDVTD-----SLNIEAPYHQASFI 491
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 458 KRGEIDFISVNLHRSR- LHWDDAEKFNPERWLPDGNPNETNQNFSLYFPFG-----GG 510
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 492 PRACFGMRFAVEMKAFILFVTLRRVQFE 519
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 511 PRKICGDMFASFNVAIMLIRRFNFQ 538
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
Q93VK5 PRELIMINARY; PRT; 595 AA.
ID Q93VK5
AC Q93VK5;
```

206	QY	206	ALAGFDYKSDSLQNTKNTNELYAFVGLTGDGAPTLDLDFKAINMWDVFPVFTMTKRRHEIPTL	266
256	Db	256	GRAVENYFDLSTNDTGVIEAVTVTLREAEKRSVPID-VWD-IPIWKDIS-----PRQ	307
266	QY	266	QGLAVSRVGVTELMKQKQAVLGASDAQVKKDQV-----GRD--ILSLVRAVIAA	316
308	Db	308	RKVATSLKINDTLD-----LIATCKRWVEEELQFHEEYVNERDPSILHFL-----	355
317	QY	317	NLPESQKSLDDEVLAAQISNLLFAGYTSSTVLTWPHRLSDKAVQDKLBEICQIDTD-	375
356	Db	356	-LASGDVSSKQLRDDLMTWMLIAGHETSAVLTWTXYLLTTTPSVAKQGEVDSVIGDR	414
376	QY	376	MPTLDLALPYLFAFKESLRDPDPSPYANERECUKDBDFIPLAEPVIGRDSGVINEVRI	435
415	Db	415	FPTIQDMKKLYTTRVNVESLRLVPPQVPLIRRSI-----DNDILGEYFI	459
436	QY	436	TKGTWMLPLFNINRSFIKXGDAEERFERWLEVDTD-----SLNSIEAPYGHQASFISG	491
460	Db	460	KRGEDIFISVNNLRSP-LHWDDAKFNFERPPLDGNFNETNQNFSLYLPFG-----CG	512
492	QY	492	PRACTGWRPFAVAKAFVLTVLRVQFE	519
513	Db	513	PKRCIGDMFASGFENVAIAMLRFRNFQ	540
RESULT 4				
037689				
ID	O97689	PRELIMINARY;	PRF;	503 AA.
AC	O97689			
DT	01-MAY-1999	(trEMBLrel. 10, Created)		
DT	01-MAY-1999	(trEMBLrel. 10, Last sequence update)		
DT	01-JUN-2003	(trEMBLrel. 24, Last annotation update)		
DE	Cytochrome P450.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
OR	[1]			
RP	SEQUENCE FROM N.A.			
RT	Schuetz E., Andag R., Wieland E., Oellerich M.;			
RT	"Porcine Cytochrome P450 with high protein homology to human CYP3A4			
RT	(Nifedipine oxidase J04449).";			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.			
CC	EMBL: AF109068; RAD04628.1; -.			
DR	GO; GO:0004497; F-monooxygenase activity; IEA.			
DR	GO; GO:0014712; F-oxidoreductase activity, acting on paired d. . .; IEA.			
DR	GO; GO:0006118; F-electron transport; IEA.			
DR	InterPro; IPR001128; Cytochrome P450.			
DR	InterPro; IPR008072; EP450_CYP3A.			
DR	Pfam; PF00067; p450; 1.			
DR	PRINTS; PR01689; EP450IICYP3A.			
DR	PRINTS; PR00385; P450.			
DR	PROSITE; PS00086; CYTOCHROME P450; 1.			
KW	Heme; Monooxygenase; Oxidoreductase.			
SEQ	SEQUENCE 503 AA; 57267 MW; 22DB19F43C5988BB CRC64;			
Query Match 14.6%; Score 417; DB 6; Length 503;				
Best Local Similarity 27.9%; Pred. No. 4.4e-22;				
Matches 160; Conservative 82; Mismatches 211; Indels 120; Gaps 24				
QY	12	GLAAPSWSAIIA--FFSYLALPRRS-----SLYNLQGNPHNTNYFTGNFLDLSARTGBEH--	63	
Db	6	GFSTETWLLATSLVLLVLYGTYSHGLFKKLIGPGRPLPYF-GN--ILGYRKGVDHFD	61	
QY	64	AKYREKYSTLRPAGIAGAPVLSNTDPKVFNNHVW-KEAYDY-----PKPGWAARVLRL	114	
Db	62	KICFQYQGNMGVFD-GRQPVLAITDDMIKVLVKECYSVFTNRRSGFRGNRT----	116	
QY	115	IATGDGVWTAGEAHKRRIMIPISLSAQVKSVMVPIFLEKGMELVDKMWDAEAKOMAV	174	
Db	117	-----ALSLADEBWKQIRITLLSPFTSGKLKEMPIIISHYGDLLVGNLFAKE-AEKG--	167	

QY 175 GESAGEKATRETEGVDVVDKWDYGRATLDWALAGADYKSDSL-----QNKTNELV 226  
 Db 168 -----KPTMKDIFGAYSMVDVITSTAGVNTDFLNNPQDPFFVENSCKLLKF 213  
 QY 227 AFVGLTDFGAPTLDSPKAIKMDVFPVPRTRKRHEIPLTCGLAVS-----RRVGIELMEOKK 283  
 Db 214 SF-----FSLP-----FUSII-----TPILEVLNVTLFPKSVNVFFMSIK 254  
 QY 284 QAVLGASDAQVDKQVQGRDILSLVVRANIAANLPESQKLSDEEVLQIQLNLLFAGYET 343  
 Db 255 R-----MKESRLKDKQHRVDFQLQMINSONSKETDTHKGLSDEELVAQGVFFIAGYET 309  
 QY 344 SSTVLTMFMHLSDEKAVQKLEETCQI---DIDMETLDELNALPYLFAFYKESIRLDP 401  
 Db 310 TSSLSLVLVELATHPDVQKQLEEDATFPKALPSYDALAQMEYLDKVNNEILRLPY 369  
 QY 402 SPYANRECLDDEFIPLAEFVIGRDSVINEVRITKTMVLMPLFNINRSKFIYGEDAAE 461  
 Db 370 AARLERVCKKQVE-----IHGVSVPKGTVMVVPFVSIHRDPDLWPE-PEE 413  
 QY 462 FRPERWLEDVTDLSNLSIEAPYHQAFISQPRACFGWRFAVEMKAPLFVTLRRVQREPI 521  
 Db 414 FRPERFSKKNKDSIN-----PYT-LFPGTGRNCIGRFPALNMKALVRVQNPSPKPC 468  
 QY 522 -----ISHPEYEHITLIISPRIVG 541  
 Db 469 KETOTPLKLSQGLIQEKBILLKWPVDRGTVG 501  
 RESULT 5  
 ID QBLIR5 PRELIMINARY; PRT; 517 AA.  
 AC QBLIR5  
 DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Putative cytochrome P450.  
 GN QJ1332.C12.12.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriactoidae; Oryzaeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=sv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 clone:OJ1332.C12.12";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AP003752; BAC10039.1; -;  
 DR Gramene; QBLIR5; -;  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; Cytochrome\_P450; 1.  
 DR Heme; Monooxygenase; Oxidoreductase.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 517 AA; 58038 MW; 3C3E35745026E138 CRC64;  
 Query Match 14.68; Score 417; DB 10; Length 517;  
 Best Local Similarity 26.58; Pred. No. 4.6e-22;  
 Matches 154; Conservative 95; Mismatches 202; Indels 130; Gaps 24;  
 QY 6 LLTGALGAAPSWASIAFFSLYIAP-----RRSSLYNQGNHNTFTGNFLDLSARTGE 61  
 Db 7 MVAANAASVAFDAVAVKLVNRPRAITRLRAQGVGGPGY-RFFSGMLGEI--RLRD 63  
 QY 62 EHA-----KYREKYSTLRPAGIAGAPVLNASTPKPVNHWMEAY 101  
 Db 64 EGAGVLDVSSHDVPLVQHPFRKRWPLYGKTPMY-WFGARPTICLADVSMVRQVLSRT 122

QY 102 D-YPK---PGMAARVLRIATGDGVTATGEAHRHRRIMIPSLSAQAVKSMVPIELEKGM 157  
 Db 133 GYPTQVSNPYFARLL-----GKGLVLTGDGEWKHKRVVHPAFNWDKLMWT----- 170  
 QY 158 ELVDKMMEDAAKMDMAGVESAGEKKATRLTEGVDVQKMWGR-----ATLDVMAAGF--D 211  
 Db 171 -----VTMSDCAQSMISEWES-----ELGTGDIVIEISLRRFELTADVISHTAFSS 219  
 QY 212 YKSD-----SLQNKTNELVAVFGLTDFGAPTLDSPKAIKMDVFPVPRTRKRHEIP 263  
 Db 220 YKEGQVFLAQRELO-----FLAF-----STLSI-----QIP 247  
 QY 264 LTQGLAVRRVGVHELMQKQAVLGASDAQVDKDV---QGRDILSLVVRANIAANLPES 321  
 Db 248 GSSVLPKTKKLLKTVSDVKVRSMLTDIISRNNKNDVAGYGNLGLMLEA-CAPEHGES 306  
 QY 322 Q-KLSDEEVLQAINLLFAGYETSTVLTMFMHRLSEDKAVQKLEETC-QIDTDMPTL 379  
 Db 307 QPOLSMDEIIAECKTFFFGAGHDTTSHLLTWTMFLSTHPEWQEKLEEVATECDGKVPTG 366  
 QY 380 DELNALPYLFAFYKESIRLDPSPYANRECLDDEFIPLAEFVIGRDSVINEVRITKGT 439  
 Db 367 DMLAKLKLVMNMLETLURLYGPVAFIQRVNALE-----LGGITVPEGT 411  
 QY 440 VMMLPLFNINRSKFIYGEDAAEERPERWLEDVTDLSNLSIEAPYHQAFISQPRACFGWR 499  
 Db 412 VLSIPATIHRRKEVWGEDADIFKPERFKNGVSKAGYPNA---LLSFSSSPRACIGCN 467  
 QY 500 FAVAEMKAPLFVTLRRVQREPIISHPEYEHITLIISPR 538  
 Db 466 FAMIEAKAVIAMILQRFSEF---TSLPKYVHVPTDVTITLPPK 505  
 RESULT 6  
 ID Q95L35 PRELIMINARY; PRT; 503 AA.  
 AC Q95L35  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Cytochrome P450 3A.  
 GN CYP3A.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Soucek P., Zuber R., Anzenbacherova E., Linka M., Anzenbacher P.;  
 RT "cDNA of minipig cytochrome P450 3A";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF242780; AAL13316.1; -;  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006112; P:oxidoreductase activity, acting on paired d. . .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01689; EP450IICYP3A.  
 DR PROSITE; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 503 AA; 57085 MW; 5A17128CA9E50914 CRC64;  
 Query Match 14.18; Score 403.5; DB 6; Length 503;  
 Best Local Similarity 27.88; Pred. No. 4.4e-21;  
 Matches 152; Conservative 81; Mismatches 191; Indels 123; Gaps 22;  
 QY 12 GLAAPSASTA--PFSLYLAPRES-----SLYLNQGNHNTFTGNFLDLSARTGE-- 63  
 Db 6 GFSTWLLATSLVLDLYGTYSHGFLFKKLIGPGRPLPYF-GN---ILGVRKGVDFD 61



```

QY 64 AKYREKYSTLRPAQI--AGAPVLNSTDPKVFNVHM-KEAY-----DYPKPGMAARVL 113
Db 62 KXCFQYQY---KMGVYDGRQLAATVTPNNIKSVLKYECYSVPTNRSFGPLGAMENAL 118
QY 114 RATGQGVVTAAGEAHKRRHRRIMISLSAQAVKSWVIFLEKGMELVDKMDAAEKDVA 173
Db 119 SL-----ADEBEWKRIITLLSTFTSGKLKEMFPIISHYDGLLVSNLRKS-AEKG-- 167
QY 174 VGESAGEKATRLTEGVDVQKDWVGRATLDVWALAGFDYKSDSLQNK-----TNELYVAF 228
Db 168 -----KPVMTKIDFCAYSMOVIITAFGWNIDSLNPDQPFVNSKKLLK 212
QY 229 VGLTQFATLDSFAKIMWDF-----VYPF-RTWKRHRHEIPTQGLAVSRVVG 275
Db 213 FGFDFDLISLITFFPLTPIFEVLNITLFPKSSVNFPTKSVKRKESRLT----- 262
QY 276 IELMEQKQAVLGSASDAQVKKVQGRDILSLVRANIANLPESOKLSDEVLQAISN 335
Db 263 ----DQKERV-----DLLQLMINSQNSKEMDPHKSLSNEELVGPGLI 301
QY 336 LIPAGYETSTVLTWMEHRLSEDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAFVK 393
Db 302 FIPAGYETTSLSLSLALVELATHPDVQKQLEIEATFPNKAPEPTYDALAQMEYLDVYV 361
QY 394 ESLRDPDPPSYANRECLDDEDFIPLAEPVIGRDSGVINEVRITKGTVMWMLPLFNINRSKF 453
Db 362 ETLRLVPIAARLERACKDVE-----IHGVFVPRKGTVVVVPVFLHRDPD 406
QY 454 IYGEDAEERPREWLEDVTDLSNLSIEAPYGHQASFISSGRACFGWRFAVAEMKAPLFVTL 513
Db 407 LMPE-PEEERPERFSKKNKDTIN-----BYTY-LPFGTGPANCIGMRPALNMWKLALVKVL 460
QY 514 RVQFEP 520
Db 461 QNFSEKP 467

RESULT 7
Q9LUD2 ID Q9LUD2 PRELIMINARY; PRT; 515 AA.
AC Q9LUD2;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Cytochrome P450 (AT3G14620/MIE1_12).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones.";
RL DNA Res. 7:131-135 (2000).
RN [3]
RN SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Tshida J., Jiang P.X., Jones T., Kaniya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.D., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,

```

```

RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB023038; BAB02394.1; -.
DR EMBL; AY022208; AAK97679.1; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 515 AA; 58643 MW; 9F4EFCF7686F55A1 CRC64;

Query Match 14.1%; Score 403; DB 10; Length 515;
Best Local Similarity 26.5%; Pred. NO. 5e-21;
Matches 134; Conservative 87; Mismatches 191; Indels 94; Gaps 19;

QY 42 NHTNYETGNFLDILSARTGESEHAKYREKYGSTLRPAQIAGAPVLNSTDPKVFNVHKEAY 101
Db 76 NLTDTHRVMPLIQ-QTVKHGK-----TSTMMGPIASVIV---TKPEHDKDLNRY 136
QY 102 DYPKPGMAARLRIATGQGVVTAAGEAHKRRHRRIMISLSAQAVKSWVIFLEKGMELVD 161
Db 127 DFPKPVHPZIVELFAT--GVALYGEKMSKHKRIINPSFLEKLIKIMIPAFYBSCSEMIS 184
QY 162 KMMEDAAEKDVAAGSAGSEKKAATRLTEGVDVQKDWVGRATLDVWALAGF--DYKSD--- 215
Db 185 KW-----EKLVTQSSNE-----IDWYLGDLTSDVISRTAFGSSVEEGKRIF 229
QY 216 SLQNKTNELYAVFGLTDFGAPTLDLSFKAIMWDFVYFTMKRHEIPTQGLAVSRVVG 275
Db 230 ELQEQCRRLV-----KALELAFIPGRFLPTKNNLRMQ---INKEVK 270
QY 276 IELME--QKKQAVLGSASDAQVKKVQGRDILSLVRANIANLPESOKLSDEVLQAQI 333
Db 271 SRLREIIMKQ-----RGMDTGEAPNDLLILLESNG-----DHGMSIEDVVVEEC 317
QY 334 SNLLFAGYETSTVLTWMEHRLSEDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAF 391
Db 318 RLFPAGQETAVLLVWTIMLSHHQKWDQARBEILKVKGNKNKPNFDSLSRLKTMMSI 377
QY 392 VKESLELDPPSYANRECLDDEDFIPLAEPVIGRDSGVINEVRITKGTVMWMLPLFNINRS 451
Db 378 LNEVLRLYPGILLGRTVEK-----ETKLGED-----MILPGQAQVVIPLWVHRD 423
QY 452 KFTYGEDAEERPREWLEDVTDLSNLSIEAPYGHQASFISSGRACFGWRFAVAEMKAF 508
Db 424 PELWGEDVHEFNPFRFADGISKATK-----NQVSFLPFGWGPFCPCQGNFALMEAKMA 476
QY 509 LFTVLRVQFE--PIISHPEYEHITL 532
Db 477 LVLLQRFSFELSPSYTHAPTVLTL 502

RESULT 8
Q8W1E1 ID Q8W1E1 PRELIMINARY; PRT; 515 AA.
AC Q8W1E1;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE AT3G14620/MIE1_12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Banh J., Bowser L.,

```

NCBI\_TaxID=3702;

(1)

SEQUENCE FROM N.A.

Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J., Bowser L., Carninci P., Chung M.K., Goldemith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin X., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

"Arabidopsis cDNA clones";

Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.

(2)

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

Rouley-S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M., Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;

Submitted (MAR-1998) to the EMBL/GenBank/DBSJ databases.

(3)

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

Lin X.;

Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.

(4)

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

Town C.D., Kaul S.;

Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.

(5)

SEQUENCE FROM N.A.

Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J., Bowser L., Carninci P., Chang E., Dale J.M., Goldemith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

"Arabidopsis ORF clones";

Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.

-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

EMBL: AF367329; AK32916.1; -

EMBL: AC004411; AAC34227.2; -

EMBL: AY091688; AAM10287.1; -

PIR: T02191; T02191.

GO: G0004497; Fimonoxygenase activity; IEA.

GO: G0006118; Pteleon transport; IEA.

InterPro: IPR001128; Cytochrome\_P450.

Pfam: PF00067; p450; 1.

PRINTS: PR00385; P450.

PROSITE: PS00086; Cytochrome P450; 1.

Heme: Monooxygenase; Oxidoreductase.

SEQUENCE 519 AA; 59603 MW; 33053C83F917457A CRC64;

Query Match 14.0%; Score 400.5; DB 10; Length 519;

Best Local Similarity 23.5%; Pred.No. 7.7e-21;

Matches 138; Conservative 110; Mismatches 231; Indels 109; Gaps 20

QY 1 MFVILLTGALGAFAFWSIAFSLAPRESSLYNLQGNHNTYTGNF----- 51

9 LIVLIL---IGLRFKAFMLVWHFVTLRLKXNQIGSGPNY-RIFYGNLSLKKWRE 64

QY 52 --LDILSARTGE-----HAKYREKYGSTLRFAGIAGAPVLNSTPKPVFNHWKREAYD 102

65 SHLSILDPSSNDIPRLPHFYQKWMQYGETLYWN-GTEPRICISDPELAKTMLS NKLK 123

QY 103 -YKPGMAARVURIATGGVVTAGEAHKRRIRIMPSLSAQAKSVVPFLSEKGMELVD 161

124 FFYKSKARPEAVKLVGSKGLVFLEGADWVRRHRLINPAFSLDRUKIMTVWVDTL----- 179

QY 162 KMWEDAAEKDMAVGESAGEKATRETEGVVDVKDWVG----RATLDVWALAGFDYKSDSL 217

180 KMLEE-----WRKSTKSTHPKPKKENBEFORLTADIATSAFG----- 221

```

QY 218 QNKTNELYVAVGLTGDGAPATLDSFKALMMDVFPYFRMKRRHEIPLTQGLAVSRVVGIE 277
Db 222 -----SSVGEIEVFRSQM-ELKRCYTTLSNQVSIPGTQYLPFPSNIRW 265
QY 278 LMEQKQKQAVLSASDAQVNDKXOVGRDILSLIVANLANLPSQKLSDEEVLQISNLL 337
Db 266 KLERQNDNSIKIISRLQSKSDYGDLLGILLKAYNTEG--KERRKWSIEEIIHCRTFF 323
QY 338 FAGYTSSTVLTMWPHRLSEDKAVQDKLREI---CQIDTDMPTLDELNALPYLEAFVKE 394
Db 324 FGCHETTNNLAWTTMLLSLHQDQWQKLEBIFKECGKE-KTPDSETFESKLLMNVMINE 382
QY 395 SRLDPPSPFANRECLDKEDFIPLAEFVIGRDSV---INEVRITKGTWMLPLFNINES 451
Db 383 SRLVGP-----VSALAREASVNIKIGDLIPKGTIVVPLLNKSHD 424
QY 452 KFIYGEDAEFPPEWLESDVDSLSNIEAPYGHQASFISGPRACGWRFAVAKMAFLV 511
Db 425 KTLWGSADAKFNMFANGVSRANHPNA---LLAFSVGPRACIGQNFVMEAKTILTM 480
QY 512 TLRRVQFELIISHPEYEH--ITLISRPVIGREKEGYQMRLOVXV 557
Db 481 ILQRFRTSLCD--EYKTPVDNVTIQD-----YGLFVMEQZEE 518

RESULT 10
Q8LHV0 PRELIMINARY; PRT; 519 AA.
AC Q8LHV0;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN P0025D09.13.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cev. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone:P0025D09."
RL Submitted (OCT-2001) to the ENBL/GenBank/DBJ databases.
CC 1. SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF004264; BAC10362.1; -.
DR Gramine; Q8LHV0; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 519 AA; 58547 MW; 77E7B9BC1428FE38 CRC64;

Query Match 13.9%; Score 397.5; DB 10; Length 519;
Best Local Similarity 26.0%; Pred. No. 1.3e-20;
Matches 155; Conservative 80; Mismatches 225; Indels 137; Gaps 22;

QY 10 ALGLAAP-----SWASIAFFSLYLAP-----RSSLNLYQGNHNTYFTGNFLDI 54
Db 2 ANGLLAWMVAAAAAVALASNAFVAVLWVRFAISRLRAQGVGGFGY-RPFSGNLGI 60

QY 55 LSGAR-----TGSEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVH 96
Db 61 KRFRGDGAGVLNVSSHDFLPVQPHFRKFIPLXGRTFLY-WFGAQNICLADSVVMQV 119
QY 97 MKEAYD-YPKFGMAARVLRATGGVYTAEGEAHKKRRRIMISLQAQAKSVNFIPEK 155

```

```

Db 120 LSDRTGIYPKNLTNPHFVRL-LGKGLVLTDGDEWKRHRKVVFHFAFNMDKLKMMT----- 172
QY 156 GMELVDKMMEDAAEKDMAVGESAGKATRLTEGVDVKDWGVGRATLDVNMALAGF--DYK 213
Db 173 -----MTMSDCSRSMSEWESELAAGGLVE-----IELSRFEELTADVISHTAFGSSYK 223
QY 214 SSS---LQKNTNELYVAVGLT---DGFA---PTLDSFKALMMDVFPYFRM-----KRRH 280
Db 224 EGKQVFLAQRELQFLAFSTFLTVQIPGFSYLPYTKNFKT--WSLDKKVRGMLMDIITRHH 281
QY 261 EPLATQGLAVSRVVGIELMEQKQKQAVLSASDAQVNDKXOV--OGRDILSLIVANLANL 318
Db 282 A-----NKDVAGYGNLGLMLACAPHEG 306
QY 319 PFSQKLSDEEVLQISNLLFAGYTSSTVLTMWPHRLSEDKAVQDKLREIC-QIDTDM 377
Db 307 EPCQLSMDDEIIDECKTFFAGHDTTSHLLTWTMFLSTHPDWQEKLEELIEAMECQDKVP 366
QY 378 TLDELNALPYLEAFVKESSLRLDPPSPFANRECLDKEDFIPLAEFVIGRDSVINEVRITK 437
Db 367 TGDMLNKLKXNNMFELETILRLYSPVSLIRKVDVTDIE-----LGGIKMPE 411
QY 438 GTWMLPLFNINRSKFIYGEDAEFPPEWLESDVDSLSNIEAPYGHQASFISGPRACFG 497
Db 412 GALLTIPTATIHRRKEVWGEDADEFRFERFENGVTAAKHNA-----LUSFSGSPKSCIG 467
QY 498 WRFVAEMKAFLVTLRRVQF--EPIISHPEYEHITLISRPVIGREKEGYQMRLO 552
Db 468 QNFAMIEAKAVIAMILQRFSTLSPKXVHAPTDVITL-----RPKYGLPLMILK 515

RESULT 11
Q86SK3 PRELIMINARY; PRT; 503 AA.
AC Q86SK3;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Cytochrome P450.
GN CYP3A43/CYP3A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21839017; PubMed=11726664;
RA Finta C., Zaphiropoulos P.G.;
RT "Intergenic mRNA molecules resulting from trans-splicing."
RL J. Biol. Chem. 277:5882-5890(2002).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Zaphiropoulos P.G.;
RL Submitted (MAY-2003) to the ENBL/GenBank/DBJ databases.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 503 AA; 57394 MW; 4141F95B28B34A6E CRC64;

Query Match 13.8%; Score 394; DB 4; Length 503;
Best Local Similarity 28.7%; Pred. No. 2.2e-20;
Matches 142; Conservative 83; Mismatches 177; Indels 92; Gaps 20;

QY 47 FTGNFLDIISARTG-----EEHAKYREKYGSTLRFAIAGAPVLNSTDPKVFNVH--KE 99

```

Db 46 FLGN---ILSHKGFMPDMECHKKYGKNGP---VDG---QQPVLAITDPDMIKTVLWKE 97  
 QY 100 AYDY---PKGMAARVLRIATGSGVVTAEAGEAHRHRRIMIPSLSAQAQVKMVFIFLEKG 156  
 Db 98 CYSVFTNRPRPGPGVGFKSA---ISIAEDBEWKRLRSLLSPTFTSGKLKEMVPIAQYG 153  
 QY 157 MELVDKMMEDAAEKDMAGVAGESAGEKKATRLTET-EGVDVKDWVGRATLDVMALAGFDYKSD 215  
 Db 154 DVLVRNLREA-----ETGKPTVLKDVFGAYSMOVITSTSEGVNID 194  
 QY 216 SLQKNTNELYVAFVGLTDFGAPILDSPKAIM-WDFV-PYRTWK-RRHEIPLTQGLAV-- 270  
 Db 195 SLNPNQD-----PFVNTKLLRDFLDPPFLGTFPPFLIPLEVLNICY 240  
 QY 271 -SRRVGIEMEQKQAVLGSASDAQVQKQVQGRDILSLVRANIAANLPSOKLSDEV 329  
 Db 241 FPREVTFKRSVKR-----MKESRLEDTQKRVDFLQLMDSQSKETESHKALSDEL 295  
 QY 330 LAQISNLLFAGYETTSVLTWTFHRLSEDAKQVODKLEIEICQI--DTMPTLDELNALPY 387  
 Db 296 VAQSTIFIFAGYETTSVLTWTFHRLSEDAKQVODKLEIEICQI--DTMPTLDELNALPY 387  
 QY 388 LEAFVKEISRLDPPSPVANRECKLDEDFIPLAEFVIGRDSVINEVRITKGMVMLPLFN 447  
 Db 356 LDMVNETLRLFPAMRLERVKCKVE-----INGMFIKGVVWIPSYA 400  
 QY 448 INRSKFYIGDEAEPRFRMLEDVTDLSNS-IEAPYGHQASISGPRACFWRFVAVEMK 506  
 Db 401 LHRDP-KYWTPEKFLPERFSKXKKNIDPIYITPFG-----SGPRNCIGMRFPALNMK 453  
 QY 507 AFLVTLRRVQFEP 520  
 Db 454 LALIRVLQNFSEKP 467

## RESULT 12

Q9FEEL PRELIMINARY; PRT; 544 AA.  
 AC Q9FEEL1;  
 DT 01-MAR-2001 (TremBLrel. 16, Created)  
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)  
 DE Putative cytochrome P450.  
 GN P068A04.9 OR P0006C01.24.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryzae; Oryza.  
 CX NCBI\_TaxID=4530;  
 RN [1]\_  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P068A04.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0006C01.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF002839; BAB19103.1; -  
 DR EMBL; AF002744; BAB19082.1; -  
 DR Gramene; Q9FEEL1; -  
 DR GO; GO:000449; P:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; p450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.

KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 544 AA; 60867 MW; 273EAF5968D1A024 CRC64;  
 Query Match 13.8%; Score 393.5; DB 10; Length 544;  
 Best Local Similarity 26.4%; Pred. No. 2.7e-20;  
 Matches 159; Conservative 83; Mismatches 215; Indels 145; Gaps 25;  
 QY 6 LLTGALGLAFAFWASIAFFSLYLAPRR-SSLYNLQGNHTNY-----FTGNFLDILSAR 58  
 Db 31 LLGNAVALLLVNAQMLEWALAPRMEPALBAQGLRGTYQYRFLHGLDTEDLRLVTAAR 90  
 QY 59 TG-----EBHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFHV 96  
 Db 91 SKPQMDRPHDFIPRVAPLHLRALEHGR-----VSFTWFGMPRVITTDPLVREV 142  
 QY 97 MKEAV-DYPKGMAARVLRIATGSGVVTAEAGEAHRHRRIMIPSLSAQAQVKMVFIFLEK 155  
 Db 143 LSNKFGHEPKTKLATRLSKLLVG-GLVILHGEKWKVHRIMNPAFHAELKEMLPFASAS 201  
 QY 156 GMLVDKMMEDAAEKDMAGVAGESAGEKKATRLTETEGVDVKDWVGRATL--DVMALAGFDYK 213  
 Db 202 CSELTGRW-----ENAVAASVQKAE-----LDI--WPDFQNLSGDVISRBAFGVR 244  
 QY 214 SDS-----LQKNTNELYVAFVGLTDFGAPILDSPKAIMMDFVPVF-----RTMKRRHEI 262  
 Db 245 HHEGRQIFLLQAQEARLV-----QSFRS---NVIPGLSYENNRMK----- 283  
 QY 263 PLTQGLAVSRV-----GIELMEQKQAVLGSASDAQVQKQVQGRDILSLVRANIAAN 317  
 Db 284 -----AIDREIKSILRGI-IEKQKATKNGAS-----KD-----DLGLLLQGNMYY 326  
 QY 318 LPESOKLSD---BEVLAQISNLLFAGYETTSVLTWTFHRLSEDAKQVODKLEIEICQI- 372  
 Db 327 SDEGSKSGMTVEEIIDECKLFYFAGMETTAVLLTWTWVALSMHPDWQDAREBILQVF 386  
 QY 373 DTDMPITDELNALPYLEAFVKEISRLDPPSPVANRECKLDEDFIPLAEFVIGRDSVINE 432  
 Db 387 GRNKPDINGVSRLLKVTWTVLHLEVRLLPYVVMNRRTYKEIE-----LGG 431  
 QY 433 VRITKGTVMVLMPLFNINRSKFYIGDEAEPRFRMLEDVTDLSNSIEAPYGHQASFIISGP 492  
 Db 432 VRYPAQVMSLPVLFIHRDAAAAGHDAGEPDPGRFAEGVARACKDPGA--GAFFPFSWGP 489  
 QY 493 RACFWRFVAVEMKAFIFVTLRRVQEP--PLIHPHYEHTLIISRPRIVGKEGKQWR 550  
 Db 490 RICIQNFALLEAKVAGLMILQRFAPFELSPAYAHAPYTVITL-----HPQHGVPVR 540  
 QY 551 LQ 552  
 Db 541 LR 542

## RESULT 13

Q8CUF2 PRELIMINARY; PRT; 497 AA.  
 AC Q8CUF2;  
 DT 01-MAR-2003 (TremBLrel. 23, Created)  
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
 DE Cytochrome P450 3A.  
 GN CYP3A.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognath.; Muridae; Rattus.  
 CX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Matsubara T., Nagata K., Yamazoe Y.;  
 RT "Isolation and characterization of a novel rat CYP3A form.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB084894; EAC23085.1; -  
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.

DR GO: 0006118; P: electron transport; IEA.  
 DR InterPro: IPR001128; Cytochrome P450.  
 DR InterPro: IPR008072; EP450\_CYP3A.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PRO1689; EP450IICYP3A.  
 DR PRINTS: PRO0385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 497 AA; 57009 MW; 8C034C25C476F452 CRC64;

Query Match 13.7%; Score 392; DB 11; Length 497;  
 Best Local Similarity 26.4%; Pred. No. 3e-20;  
 Matches 140; Conservative 81; Mismatches 176; Indels 134; Gaps 20;

QY 35 LYNLQPNHTNY-----PTGNFLDILSARTG-----BEHAKYREKYGSTLRPAG 78  
 DB 22 LLYGTSTHGNFKKLGISGPKLPFVGN---ILAYRGFWFDRHC--HKYKIDWGFYE 76  
 QY 79 IAGAPVNSTDPKVFHYM--KEAY-----DYPKPGAAARVRIATGDDGVVTAEGEAKH 130  
 DB 77 -GRQPLAITDPDIITKVLVKECYSTFNRRSFGPAGILKAITL-----SEDEWK 127  
 QY 131 RHRRIMPISLQAQVSMVPIFEKGMELVDKXMDAEDKMDAVGESAGEKKATRLTEG 190  
 DB 128 RUTLLSPFTSGKLEHFPPI-INQADLLVKNVKEAEG-----NP 169  
 QY 191 VDVKDVGVRATLDVNALAGFDYKSDSLQNTN-----ELYVAPVGLTDF 235  
 DB 170 ITMKDIFGAYSDVITGTSFGVNVDSLNNPQNPVQVKKLLKFNFLDPFLSVLPFEL 229  
 QY 236 APTLDSFKAIM--DFVYFETMKRHEIPLTQGLAVSRVGLIELMEQKQAVLGASQ 293  
 DB 230 TPVEAFDITVPKVMKFFRT-----SVERMKENR-----260  
 QY 294 AVDKDVGQRDILSLVRANAAANLPSQKLSDEVLQAI-SNLLFAGYETSTSVLTMMFH 353  
 DB 261 -MQEKVKQRLDPLQMLNSQSGSKESHQGLTDEIQAQIFIFAGYETTSALSALY 319  
 QY 354 RSEDKAVODKLRREICQIDTMP-----TLDELNALPYLEAFVKESLRDLDPSPYANRE 408  
 DB 320 LLAHPDLQKQLODE--IDAALPKAVTYDVLVMEYLDVNLVLELRLFPVGGLELV 376  
 QY 409 CLKDEDFIPLAEPIVGRDGSVINEVRIKGTMMVLPFNINRSKFIYGEDAEFRERWL 468  
 DB 377 CKKDVE-----INGVFIPKGTVMVPTFALHKDPKWPB-PEEFCSERER 420  
 QY 469 EDVTSLSN-IEAPYGHQASFSIGPRACFGWRFAVAENKAFLVTLRRVQF 518  
 DB 421 KKNQDSINPIYLPFG-----NGPRNCIGRFPALMMKIALVRLVQLNPSF 465

## RESULT 14

Q8ISJ7 PRELIMINARY; PRT; 492 AA.  
 AC Q8ISJ7  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Cytochrome P450 CYP414.  
 GN CYP414.  
 OS Mamestra brassicae.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;  
 OC Noctuidae; Hadeninae; Mamestra.  
 OX NCBI\_TaxID=55057;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Antenna;  
 RX MEDLINE=21996093; PubMed=12000647;  
 RA Maibeche-Coisne M., Jacquin-Joly E., Francois M.C.,  
 RA Nagnan-Le Meillour P.;  
 RT "cDNA cloning of biotransformation enzymes belonging to the cytochrome  
 P450 family in the antennae of the noctuid moth Mamestra brassicae";  
 RL Insect Mol. Biol. 11:273-281(2002).

DR EMBL: AY063501; AAL48300.1; --  
 DR GO: 0006118; P: electron transport; IEA.  
 DR InterPro: IPR001128; Cytochrome P450.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PRO0385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 492 AA; 56219 MW; 4F69897198D9D01B CRC64;

Query Match 13.6%; Score 389; DB 5; Length 492;  
 Best Local Similarity 22.9%; Pred. No. 5e-20;  
 Matches 139; Conservative 105; Mismatches 182; Indels 180; Gaps 24;

QY 1 MFTLVLLTGALGLAAFAFASIAFSLYLAPERSLLYNLQGN-----HNYE---TGPF 51  
 DB 6 VFVVVLA-----LLVSWISLV-----RVSRFNVPGSPPLVGNNAHLFWKSEF 52  
 QY 52 LILSARTGEHAKYREKYGSTLRPAGIA-----GAPVLNSTDPKVFHYMKEA 100  
 DB 53 LNLVQ-----RLSEKYGHAFVHFSTFPYVYVICHSKYAEPVLSSTE-----HITK-- 97  
 QY 101 YDYPFEGMAARVLRITATGDGVVTAEGEAHKEHRRIMPISLQAQVSMVPIFEKGMELV 160  
 DB 98 -----GRSYSLTLCWLGGLLTATGQWKSHRFLTPAFENILQNFLVFCNQRLT 151  
 QY 161 DKM--MEDAAEKMDAVGESAGEKKATRLTEGVVDKDWGRATLDVNALAGFDYKSDSL- 217  
 DB 152 EKIRGNADGRPIDM-----FPIIALAALDNVNESIM 182  
 QY 218 -----QNKTNELYVAFVGLTDFAPTLDSFKAIMWDVFPVFTMKRHEIPLTQGLAVS 271  
 DB 183 GVCMDAQKHSQSEYV-----KSIELSAIV-----TM--RMQIPFFGEDAIF 222  
 QY 272 RRVGIELMEQKQAVLGASDAQV-----DKKDVGQRD---ILSLVRA 312  
 DB 223 NLUPYKTKQKALKVYHGCTNKNVARRAELKANKANITLNDSSDITGPKNHTFLDLLLA 282  
 QY 313 NTAANLPESQKLSDEVLQAI-SNLLFAGYETSTVLTMMFHLSEOKAVODKLRBEI--- 369  
 DB 283 EI-----DGKIDDSVREEVDVTFMFEHDTTSGIVYTLHCLSKGRDVQEKIYBELKTI 337  
 QY 370 --CQDMDPTDELNALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAEPIVGRDG 427  
 DB 338 FGSEIHRD--FTYHELQOMKYJELVKSMSRLFPVPLIERIMKDCB-----383  
 QY 428 SVINEVRIKGTMMVLPFNINRSKFIYGEDAEFRERWLEDVTSLSNIEAPYGHQAS 487  
 DB 384 --VGGLKLVKGTSVVMNIFQIQROPDLF-DPLEFRPERF-----EAPLKNPFS 429  
 QY 488 ---FISGPRACFGWRFAVAENKAFLVTLRRVQFEPFIISHPEYEHITLIISRPRIYGRK 544  
 DB 430 WLAFGAGPNCIGQKFAWMLKTIISVKNVFFILPAABEPFELS-----ADLVIRSK 481  
 QY 545 EGYQMR 550  
 DB 482 NGVNVK 487

RESULT 15  
 Q8LQ77 PRELIMINARY; PRT; 523 AA.  
 AC Q8LQ77  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Putative cytochrome P450.  
 GN B10510.19.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.

[illegible]

Search completed: April 2, 2004, 14:00:40  
Job time : 41.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 11 Seconds  
(without alignments)  
2636.643 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLLTGALGAAAFNAS.....RIVGREXEGYQMLQKPV 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414	14.5	503	1 CP3T_PIG	P79401 sus scrofa
2	413	14.5	503	1 CP39_RAT	P1538 rattus norv
3	408.5	14.3	503	1 CP30_SHEEP	Q29496 ovis aries
4	404	14.2	503	1 CP3C_CANFA	P24463 canis famil
5	402	14.1	501	1 CP36_RABIT	P11707 oryctolagus
6	399	14.0	503	1 CP33_HUMAN	P05184 homo sapien
7	398	14.0	503	1 CP3D_MOUSE	Q84184 mus musculu
8	396.5	13.9	502	1 CP34_HUMAN	P08684 homo sapien
9	395.5	13.9	503	1 CP38_WACEA	P33268 macaca fasc
10	394	13.8	504	1 CP3B_MOUSE	Q54459 mus musculu
11	391	13.7	504	1 C341_MOUSE	Q9jma7 mus musculu
12	385.5	13.5	503	1 CP37_HUMAN	P24462 homo sapien
13	385	13.5	502	1 CP35_HUMAN	P20815 homo sapien
14	385	13.5	507	1 CP3S_BOVIN	P79102 bos taurus
15	377.5	13.2	503	1 CP3L_CALJA	O18993 callitrix
16	371.5	13.0	501	1 CP3V_WESAU	O70537 mesocricetu
17	370	13.0	504	1 CP31_RAT	P04800 rattus norv
18	368	12.9	504	1 CP3G_MOUSE	Q64481 mus musculu
19	362	12.7	503	1 CP3F_CAVPO	Q64406 cavia porce
20	359.5	12.6	511	1 CP4B_MOUSE	Q64462 mus musculu
21	356.5	12.5	496	1 C330_FUNHE	Q9pve8 fundulus he
22	355.5	12.5	503	1 CP3E_CAVPO	Q64417 cavia porce
23	351.5	12.3	496	1 C356_FUNHE	Q8axv5 fundulus he
24	351	12.3	520	1 CP38_HUMAN	P38187 homo sapien
25	350.5	12.3	1054	1 CYPE_BACSU	O08336 bacillus su
26	350	12.3	504	1 CP32_RAT	P05183 rattus norv
27	350	12.3	526	1 CP35_RAT	P1870 rattus norv
28	349.5	12.3	524	1 CP31_RAT	P33274 rattus norv
29	348.5	12.2	511	1 CP4B_HUMAN	P13584 homo sapien
30	346.5	12.1	511	1 CP4B_RAT	P15129 rattus norv
31	346.5	12.1	524	1 CP3C_HUMAN	Q2hcs2 homo sapien
32	345.5	12.1	520	1 YRV2_CAEEL	Q27514 caenorhabdi
33	344.5	12.1	502	1 C340_ORYLA	Q98t91 oryzias lat

34	344	12.1	520	1 CP3B_HUMAN	Q08477 homo sapien
35	344	12.1	524	1 CP3B_HUMAN	Q9hb16 homo sapien
36	340.5	11.9	524	1 CP72_CATRO	Q05047 catharanthu
37	340	11.9	503	1 CP3P_MOUSE	Q09158 mus musculu
38	338.5	11.9	503	1 C343_HUMAN	Q9hbs5 homo sapien
39	337	11.8	496	1 C4AE_DROME	Q46054 drosophila
40	337	11.8	503	1 CP3H_CAVPO	Q84409 cavia porce
41	337	11.8	520	1 CP32_HUMAN	P78329 homo sapien
42	336.5	11.8	518	1 CP3R_ONCMY	O42563 oncothyachu
43	336	11.8	522	1 CP34_RAT	P51869 rattus norv
44	334.5	11.7	509	1 YRV4_CAEEL	Q27516 caenorhabdi
45	334.5	11.7	580	1 C973_ARATH	O23365 arabidopsis

#### ALIGNMENTS

##### RESULT 1

CP3T\_PIG  
ID\_CP3T\_PIG STANDARD; PRT; 503 AA.

AC P79401, 1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cytochrome P450 3A29 (EC 1.14.14.1) (CYP11A29).  
GN CYP3A29.

OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9623;  
RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Small intestine;

RX MEDLINE=98347363; PubMed=9682441;

RA Nissen P.H., Winteroe A.K., Fredholm M.;

RT "Mapping of porcine genes belonging to two different cytochrome-P450

subfamilies.";

RL Anim. Genet. 29:7-11 (1998).

CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate

monooxygenases. In liver microsomes, this enzyme is involved in an

NADPH-dependent electron transport pathway. It oxidizes a variety

of structurally unrelated compounds, including steroids, fatty

acids, and xenobiotics.

CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +

oxidized flavoprotein + H(2)O

CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.

CC -!- INDUCTION: P450 can be induced to high levels in liver and other

tissues by various foreign compounds, including drugs, pesticides,

and carcinogens.

CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Z93099; CAB07513.1; -.

DR HSSP; P14779; IJJPZ.

DR InterPro; IPR001128; Cytochrome P450.

DR InterPro; IPR008072; EP450\_CYP3A.

DR Pfam; PF00087; P450; 1.

DR PRINTS; PR01689; EP450IICYP3A.

DR PRINTS; PR00385; P450.

DR PROSITE; PS00086; CYTOCHROME\_P450; 1.

KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;

KW Microsome; Endoplasmic reticulum.

FT METAL 442 442 IFON (HEME AXIAL LIGAND) (BY SIMILARITY).

SQ SEQUENCE 503 AA; 57198 MW; E3D411B2674FD17F CRC64;

Query Match 14.5%; Score 414; DB 1; Length 503;

Best Local Similarity 27.4%; Pred. No. 8.9e-17;  
Matches 158; Conservative 85; Mismatches 197; Indels 136; Gaps 23;

QY 12 GLAFAWASIA--PFSLVAPRRS-----SLYNLQGNHNTWYFTGNFLDLSARTGEH-- 63  
Db 6 GSTETWLLATSLVLLLYCTYSHGLFKLGTGPRPLPYF-GN---ILYRKGVDFD 61  
QY 64 AKYREKYGSTURFAGI--AGAPVLNSTDPKVFHVM-KRAY-----DYPKPGMAARVL 113  
Db 62 KCFQOYQ--KMGVYDGRQFLAVTDPNMKSILVKCEYSVFTNRRSFGPLGARVAL 118  
QY 114 RATGDGVVTAAGGAHKKHRRIM:PSLSAQAVKSWVP:PLEKGMELVDQWMDAAEKDWA 173  
Db 119 SL-----AEDBEWKEIRLLSPTFTSGKLEMPILSHYGDLLVSNLRE-REKG-- 167  
QY 174 VGESAGEKATRLTEGVVDVQVGRATLDYMLAGFDYKSDSLQNK-----TNELYVAF 228  
Db 168 -----KPVTKMDIFGAYSMDVITSTAFGVNIDSLNNPQDPFFVENSKKLLK 212  
QY 229 VGLTGFAPTLDSPKAIMWDF-----VPYF-RTMKRHEIPLTQGLAVSRVVG 275  
Db 213 SFSPDPFLLSLFPPLTPIPEVLNITLFPKSSWNFTKSVKEMKESRLT----- 262  
QY 276 IELMQKQAVLGASDAQVDKQVQGRDILSLVLRANTAAANLPESQKLSDEVLQAQSN 335  
Db 263 -----DQKKRV-----DLLQLMNSQNSKENDPHKLSNBEELVAQGI 301  
QY 336 LUFAGYETSSTVLTWFMHRLSEDAVDKLRRETCOI--DTDMPTLDELNALPYLEAFVK 393  
Db 302 FIFAGYETTSALSILAVELATHPDVQKQOEIATEFPKAPTVDALQAQVEYLDVNVN 361  
QY 394 ESLRLDPSPVANRECLADEFIPLAEPVIGRGSVINEVRITKGTWMLPFLFNINRKF 453  
Db 362 ETLRLYPTAARLEACKQDVE-----IHGVFVPGKVVVVVVFVLRDPD 406  
QY 454 IYGEDAEFRPERMLEDVTDLSNIEAPYGHQASFSISGRACGMRFAVAENKALFLVTL 513  
Db 407 LWPE-PEEFRPERESKHKDTIN---PYTY-LPFGTGPENCIGMRFAVMNKLALVRVL 460  
QY 514 RVQVEPI-----ISHPEYEHITLISR 536  
Db 461 QNFSRPEKQETQIPUKLTQGLTOPEKPVVILKILPR 496

RESULT 2  
CP39 RAT STANDARD; PRT; 503 AA.  
AC P51578; Q64557; Q64631;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Cytochrome P450 3A9 (EC 1.14.14.1) (CYP11A9) (P450-OLF3) (Olfactive) (3AH15).  
DE CYP3A9.  
GN Rattus norvegicus (Rat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=96220175; PubMed=8990268;  
RA Mahne A., Strotkamp D., Roos P.H., Hanstein W.G., Chabot G.G.,  
RA Nef P.;  
RT "Expression and inducibility of cytochrome P450 3A9 (CYP3A9) and  
RT other members of the CYP3A subfamily in rat liver."  
RL Arch. Biochem. Biophys. 337:62-68 (1997).  
RL [2]  
RP SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RC MEDLINE=96220175; PubMed=8660328;  
RX Wang H., Kawashima H., Strobel H.W.;  
RA "CDNA cloning of a novel CYP3A from rat brain.";

RL Biochem. Biophys. Res. Commun. 221:157-162 (1996).  
CC -!- FUNCTION: This isozyme seems to be implicated in oifaction. Active  
CC in the demethylation of erythromycin as well as benzphetamine.  
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
CC oxidized flavoprotein + H(2)O.  
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -!- TISSUE SPECIFICITY: Mainly expressed in olfactory epithelium.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U60085; AAB03662.1; -;  
CC EMBL; U46118; AAC52582.1; -;  
CC PIR; JCA702; JCA702.  
CC HSP; P14779; 1JPZ.  
CC InterPro: IPR001128; Cytochrome P450.  
CC InterPro: IPR008072; EP450\_CYP3A.  
CC Pfam; PF00667; P450; 1.  
CC PRINTS; PR01689; EP450; 1.  
CC PROSITE; PS00886; CYTOCHROME P450; 1.  
CC KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Olfaction.  
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
FT CONFLICT 457 457 F -> V (IN REF. 2).  
SQ SEQUENCE 503 AA; 57811 MW; 9BBB13E6908758B4 CRC64;  
  
Query Match 14.5%; Score 413; DB 1; Length 503;  
Best Local Similarity 27.4%; Pred. No. 1e-16;  
Matches 145; Conservative 86; Mismatches 178; Indels 120; Gaps 22;

QY 54 ILSARTG-EHAKY-REKYGSTLRFAGI--AGAPVLNSTDPKVFHVM-KEAY----- 101  
Db 50 ILAYKGFWEFDKCHKYK--KLWGLYDGRQVLAITDPIIKTVLKYCYSTFTNRR 106  
QY 102 DYPKGMARVLRATGDGVVTAAGSAHKKHRRIM:PSLSAQAVKSWVP:PLEKGMELVD 161  
Db 107 NFGPVGILKXAI-----SEDEWKRALLSFTTSGKLEMPFI----- 149  
QY 162 KWMEDAEEKOMAVGESAGEKATRLTEGVVDVQVGRATLDYMLAGFDYKSDSLQNK 221  
Db 150 NQYTDMLVRNMROGSEKGP-----TSMKDIAGYSMDVITATSGVNVDSLNNPQ 200  
QY 222 NELYVAFVGLTGDGAPTLDSFKAIM-WD-FVPFRTMKRHEI-PLTQGLAYSR----- 272  
Db 201 D-----PFVEKVKLLKFDIFDPLFLSVTLFPFLTPFEALNVSMFPRDVI 246  
QY 273 ---RVGIELMEQKQAVLGASDAQVDKQVQGRDILSLVLRANTAAANLPESQKLSDEEV 329  
Db 247 DFFKTSVERMKENR-----MKEKQKQWDFLQLMNSQNSKVKXDSHKSLDVEI 295  
QY 330 LAQISNLFPAGYETSSTVLTWFMHRLSEDAVDKLRRETCOITDMP-----TJDELNA 384  
Db 296 VAQSVIFIFAGYETTSALSILVLAHPDIQKQJDE---IDAAALPNKATATYDILLQ 352  
QY 385 LPLYEAFVKESLRLDPPSPVANRECLADEFIPLAEPVIGRGSVINEVRITKGTWMLP 444  
Db 353 MEYLDWVNETLRLVPIAGRLERVKCTDVE-----INGVFIPKGTWMP 397  
QY 445 LFNINRSFITYGEDAEFRPERMLEDVTDLSNIEAPYGHQASFSISGRACGMRFAVAE 504  
Db 398 TFLHKDHPH-YMPPEEPRPERFSKKNQDNIN---PYMY-LPFGNGPNCIGMRPALMN 451  
QY 505 MKAFLEVTILRRVQPEPI-----ISHPEYEHITLISRPRIV 540  
Db 452 MKVALFRVLQNPSPFOPCKETQIPKLKQGLQPEKPELLKVVSRDET 500



```

RESULT 3
CP30_SHEEP
ID CP30_SHEEP STANDARD; PRT; 503 AA.
AC Q29456;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A24 (EC 1.14.14.1) (CYP3A24).
GN CYP3A24.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC Ching M.S., Chun-Jing J., Ghabrial H., Wooley P.J., Smallwood R.A.,
RA Morgan D.J.;
RL "Ovine foetal liver CYP3A24."
CC Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J59378; AA02657.1; -.
DR HSP; P14779; IJPE.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450I1CYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57360 MW; 804223EAD0304238 CRC64;

Query Match 14.3%; Score 408.5; DB 1; Length 503;
Best Local Similarity 28.5%; Pred. No. 1.8e-16;
Matches 165; Conservative 82; Mismatches 197; Indels 135; Gaps 27;

QY 7 LTGALGLAFAFNASIAFFSLYLAPRRS----SLYNLQPNHTNYFTGNFLDLSARTG-- 60
DB 3 LIPFSLETWLVLLAISLVLLVLYGYSHGLFKLVGSGRPLPYF-GN--VLSRYKGV 58
QY 61 ----BEHAKYREKYGSTLRPAGIAGAPVNLSTDPKYNHVM-KEAYD-----YKPGM 108
DB 59 EFDEECFKYKQWGV---FEG--KQPLVITDPPVTKVTLVKECYSVFTNRRVFGPMGI 113
QY 109 AARVLRIATGCVTAEGEAKHRRRIIMPISLQAVKSMVPIFEKGMELVDKMEADA 168
DB 114 MK-----NAVVAEDQWKRIRITLLSPFTSGKLDMPFLIGKYGDVLRNLRKE-A 164
QY 169 EKDVAAGESAGEKKATRLTEGTGVVDKQVGRATLDWMALAGFDYKSDSLQNKINLYVAF 228

```

```

Db 165 EK---GKS-----VNMKIDFGAYSMVDITSTSGVNDISLGNP----- 199
QY 229 VGLTDFGAPT-----LDSP--KAIMWDF-VPYFR-----TMKRRHEIP-LTQGLAVS 271
DB 200 ---QDPFVENAKLLRNFNLDLPFLSVLFPFVIFVFNITMPFKSAVDFTK--SYK 254
QY 272 RRVGIELMECKQAVLGSASDAQVDKQVQDILSLVVRANIAANLPESQKLSDEEVLA 331
DB 255 RIKESRLKQKQKPRV-----DFLQMINSONSKETDNHKLASQELMA 297
QY 332 QISNLLPAGYETSTVLTMFHLRSLSDKAVQKLRREICQIDTM-----PTLDELNALP 386
DB 298 QSVIFIFAGYETTSNTLSFLLYILATHPDVQKQLEB---IDATFPNKAFTYDVLQME 354
QY 387 YLEAFVKESRLDPPSPYANRECKLDEDFIPLAEPVIGRSGVINEVRITKGTWMLPLF 446
DB 355 YLDMVWNETLUMFPIAVRLDRCKQVE-----IHGVSIPKGTAVTVPIF 399
QY 447 NINRSKFIYGEDAEFRPERLEWEDVTDLSNS-IEAPYHQASFISGFACFRGWRFAVEM 505
DB 400 VLHEDPOLWPE-PEEPRPERFSKKNKDSINPYVLPFG-----TGPRNCIGMRFAIMNM 452
QY 506 KAFLEVTLRVQVEPIIHPYEHITLISRPVIGREK 544
DB 453 KLAIVRVLQNFSPKPC---KETQIFPKINSQGLIRPEK 487

RESULT 4
CP3C_CANPA
ID CP3C_CANPA STANDARD; PRT; 503 AA.
AC P24463;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A12 (EC 1.14.14.1) (CYP3A12).
GN CYP3A12.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Beagle; TISSUE=Liver;
RC MEDLINE=91159488; Pubmed=2001406;
RA Ciaccio P.J., Graves P.E., Bourque D.P., Glimmann-Gibson B.,
RA Halpert J.R.;
RT "cDNA and deduced amino acid sequences of a dog liver cytochrome
RT P-450 of the IIIa gene subfamily."
RL Biochim. Biophys. Acta 1088:319-322(1991).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54915; CAA38687.1; -.
DR PIR; S14275; S14275.

```

[REDACTED]

```

69  M9GLFDGRQPLNVTDPDMIKTVLWKECVSVFTNRSFGVFMKKA VSI-----SE 120
126  G5AHFEHRRIMIPSL5AQA VSKVMVPIEKGEMELVPMKMDAAEKMDMAGESAGEKKATR 185
121  DEDWKRVRTLLSPFTSGUKKEMLPITIAQGDVLV-KNLRQBAEKG-----165
186  LETEGVDVDMVGRATLDMVALAGFDYKPSLONKNTNE-----LVVAFVGLTGDGPAP 237
166  ---KPVDLKEIFGAYSMVDVTGTSFGVNIDSLRNPDPPVKVNRLLKFSF-----FDP 216
238  TLDSFKALIMWDPVYFRTMKRHEIPLTQGLAVSRVRGIELMEQKQAVLGSASDAQVDK 297
217  LLTSI-----TLFPFL-----TPIFEALHIS-MFKDVMDFUKTSVEKTKDRLKDK 262
298  KVVQGRDILSLVRANTAANLPSQKLSDEVLIAQISNLLFAGYETSTSVLTMPHRLSE 357
263  QK-RRVDFQLQMLNQSQSKSDHSHKALDDIEVVAQSIILFAGYETSTLSIFMHLIAT 321
358  DKAVQDKLRREICQIDTDMP-----TLDELNALPYLEAFVKGSLRLDPPSPYANRECLD 412

```

Db 322 HPDQVKQLOBE---IDTLPNKELATYDILVKNEXYLDVNVNLTLLYPIAGRLERVCKD 378

Qy 413 EDPIPLAEPVIGRDSGVINEVRITKGTMTWMLPFLNINRSKFIYGEDAEFRPERLWEDVT 472

Db 379 VD-----INGTFIPKGTIVMPTALHRRDPQWTE-PDBERERSKNK 422

Qy 473 DSLNS-IEAPYHQASFISSPRACFGWRFAVAKMAFLVTLRRVQFE-----PI--- 521

Db 423 DNINPYIYHFFG-----AGPRNCLGRFALMNIKLALVRLMNFSFKLCKETQVPLKLG 476

Qy 522 ---ISHPEYEHITLIISRPRIV 540

Db 477 KQGLQPEKEIVLKVSRDGI 498

RESULT 6

CP33 HUMAN STANDARD; PRT; 503 AA.

AC P05184;

DT 13-AUG-1987 (Rel. 05, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome P450 3A3 (EC 1.14.14.1) (CYPIIIA3) (Hlp).

GN CyP3A3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=86259780; PubMed=3460094;

RA Molowa D.T., Schuetz E.G., Wrighton S.A., Watkins P.B., Kremers P.,

RA Mendez-Picon G., Guzelian P.S.;

RT "Complete cDNA sequence of a cytochrome P-450 inducible by

RT glucocorticoids in human liver.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:5311-5315 (1986).

RN [2]

RP SEQUENCE OF 1-20.

RC TISSUE=Liver;

RX MEDLINE=86298342; PubMed=3898085;

RA Watkins P.B., Wrighton S.A., Maurel P., Schuetz E.G., Mendez-Picon G.,

RA Parker G.A., Guzelian P.S.;

RT "Identification of an inducible form of cytochrome P-450 in human

RT liver.";

RL Proc. Natl. Acad. Sci. U.S.A. 82:6310-6314 (1985).

CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate

CC monooxygenases. In liver microsomes, this enzyme is involved in an

CC NADPH-dependent electron transport pathway. It oxidizes a variety

CC of structurally unrelated compounds, including steroids, fatty

CC acids, and xenobiotics.

CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +

CC oxidized flavoprotein + H(2)O.

CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.

CC -!- INDUCTION: By glucocorticoids

CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; D00003; BAA00001.1; -

DR EMBL; M13785; AAA35742.1; -

DR PIR; A29410; A29410.

DR HSSP; P14779; LUPZ.

DR Genew; HGNC:2636; CYP3A3.

DR GO; GO:0005792; C:mitochondrion; TAS.

DR GO; GO:0015034; F:Cytochrome P450 activity; TAS.

DR InterPro; IPR001128; Cytochrome\_P450.

DR InterPro; IPR008072; EP450\_CYP3A.

DR Pfam; PF00067; P450; 1.

DR PRINTS; PRO1689; EP450IICYP3A.

DR PRINTS; PRO3385; P450.

DR PROSITE; PS00086; CYTOCHROME\_P450; 1.

KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;

KW Microsome; Endoplasmic reticulum.

FT INIT MET 0

FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

SQ SEQUENCE 503 AA; 57428 MW; 9885D1F729658FPC0 CRC64;

Query Match 14.0%; Score 399; DB 1; Length 503;

Best Local Similarity 28.4%; Pred. No. 6.4e-16;

Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;

Qy 13 LAAPSW--ASIAFFSILYAPRES-SLYNLQG--PNTNY--FTGNFLDILSARTG-----E 61

Db 6 LAMETWLLAVSLVLLYLYGTHSHGLFVKLPGGTFPLFLGN-----ILSYHKGCFMPDME 62

Qy 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVNHYM--KEAYD-----YKPGVAAARVL 113

Db 63 CHKYGKVGPF---YDG--QQPVLAITDPDMIKLVKVECYSVFTNREFPGVGFMSAI 117

Qy 114 RIATGCVVTAEGEAKHRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDWA 173

Db 118 SI-----AEDEWKRLRLSLLPTFTSGKUKEMVPIIAQYGDVLRNLRRE----- 162

Qy 174 VGESAGEKKATRLTEGVVDKDWGRATLDVVALAGFDYKSDSLQNKTNELYVAFVGLTD 233

Db 163 -----RETGKPVTLKDVFGAYSMDVITSSSGVNVDSLNNPD----- 200

Qy 234 GPAPTLDSPKAIM-WDFV--PYFRTMK--RRHEPLTQGL-----AVSRVG 275

Db 201 ---PLVENTKKLRFDFLDFLFSITVFPFLPILEVLNVCVFPREVTNFKAVKRWKE 257

Qy 276 IELMEQKQAVLGASDAQVDKQVGRDILSLVRANIAANLPESQK--LSDEEVLQIS 334

Db 258 SRLDETQKRV-----DFLQIMTDSHKNSKETESHKALSLELVAQSI 300

Qy 335 NLLFAGYETSTVLTWMFHLRSLSDKAVQDKLREICQI--DTDMTLDELNALPYLEAFV 392

Db 301 IFIFAGYETSTVSLTFMVELATHPDVQKQLOBEIDAVLPNKPAPPTDYTLQMEYLDWV 360

Qy 393 KESLRDPPSPYANRECLKDEDFIPLAEVPIGRDGSVINEVRIKGTVMVMLPFLNINRSK 452

Db 361 NETLRDPPFAMRLERVCKDVE-----INGMFIPKGVWVMIPISYALHRDP 405

Qy 453 FIYGEDAEERPERPRLWLEDVTDLSLNS-IEAPYHQASFISSPRACFGWRFAVAKMAFLV 511

Db 406 -KYWTEPEKFLPERFSKKNKDIDPIYTFPG-----SGPRNCIGMRPALMNMKLALIR 458

Qy 512 TLRRVQFEP 520

Db 459 VLQNFSEKP 467

RESULT 7

CP3D\_MOUSE STANDARD; PRT; 503 AA.

ID CP3D\_MOUSE

AC Q64464;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome P450 3A13 (EC 1.14.14.1) (CYP3A13).

GN CYP3A13 OR CYP3A-13

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ddy; TISSUE=Liver;

RX MEDLINE=95101705; PubMed=7803471;





```

RESULT 10
CP3B MOUSE
IID CP3B MOUSE STANDARD; PRT; 504 AA.
AC Q6459;
15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome P450 3A11 (EC 1.14.14.1) (CYP1A11) (P-450I1A1) (P-450I1A1)
DE 450UT;
DE CYP3A11 OR CYP3A-11.

```

```

RESULT 9
CP38 MACFA
ID CP38 MACFA STANDARD; PRT; 503 AA.
AC P3368; P25231;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 29-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A8 (EC 1.14.14.1) (CYP3A8) (P450-MKNF2) (P-450-MK2).
GN CYP3A8.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OC NCBI_TaxID=9541;
OX [1] RN
OX [1] SEQUENCE FROM N.A.
RP RN
RP SEQUENCE FROM N.A.
RC RN
RC TISSUE=Liver;
RX MEDLINE=93129612; PubMed=1282830;
RA Komori M., Kikuchi O., Sakuma T., Funaki J., Kitada M., Kamataki T.;
RT "Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity
RT of the primary sequences to human cytochromes P-450.";
RL Biochim. Biophys. Acta 1171:141-146(1992).
RN [2]
RP SEQUENCE OF 1-22.
RC RN
RC TISSUE=Liver;
RX MEDLINE=89287352; PubMed=2500151;
RA Ohta K., Kitada M., Hashizume T., Komori M., Ohi H., Kamataki T.;
RT "Purification of cytochrome P-450 from polychlorinated biphenyl-
RT treated crab-eating monkeys: high homology to a form of human
RT cytochrome P-450.";
RL Biochim. Biophys. Acta 996:142-145(1989).
CC CC
CC -!- FUNCTION: Catalyzes nifedipine and nifedipine oxidations.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By polychlorinated biphenyls (PCB).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ddv; TISSUE=Liver;  
 RX MEDLINE=92223116; PubMed=1339292;  
 RA Yanagimoto T., Itoh S., Muller-Eberhard D., Kametaki T.;  
 RT "Mouse liver cytochrome P-450 (P-450I11A1): its cDNA cloning and  
 RT inducibility by dexamethasone.";  
 RL Biochim. Biophys. Acta 1130:329-332 (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Souffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
 CC -!- FUNCTION: Catalyzes erythromycin N-demethylation, nifedipine  
 CC oxidation and testosterone 6 beta-hydroxylation.  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- INDUCTION: By dexamethasone.  
 CC -!- TISSUE SPECIFICITY: Highly expressed in liver.  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X60452; CAA42981.1; --  
 DR EMBL; BC010528; AAH10528.1; --  
 DR PIR; S22334; A60564.  
 DR HSP; P14779; IJPF.  
 DR MGI; 88609; Cyp3a11.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01689; EP450ICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 504 AA; 57854 MW; E369AF71CE23F180 CRC64;  
 Query Match 13.8%; Score 394; DB 1; Length 504;  
 Best Local Similarity 25.6%; Pred. No. 1.2e-15;  
 Matches 144; Conservative 103; Mismatches 215; Indels 100; Gaps 23;

QY 7 LTCALGAAFSWASIAFFSLY-LAPRRSLYNLOG-----PNHTNYFTGNFLDILSAR 58  
 DB 3 LVSALSLETWLLAISLVLYRYGRKHELFKKQILPGPKLP-----FLGTVLNYKGL 57  
 QY 59 TGEHAKYREKYGSTL-REAGIAGAPVLNSTDPKVFHVM-KEAY-----DYKPKGNA 109  
 DB 58 WKFDMECYK-KYGTWGLFDG--QTPLAVTDPETIKNVLVKECESVFTNRDRDFGVGIM 114  
 QY 110 ARVLRTATGQVVTAGEAHKRRIMIPSLSAQAVKSVPIFLEKGMELVDKMMEDAAE 169  
 DB 115 SKAISISKOD-----EWKRYRALLSPTFTSGKLKMFVIEQYQ-DILVKYLRQAK 165  
 QY 170 KDMAVGESAGEKATRLTEGVDVKWVGRATLDVMAAGFDYKDSLQNKTNELYAVFV 229  
 DB 166 KG-----KPVTKMDVLGAYSMDVITSTSGVNVDSLNNPDPFVERAK 208  
 QY 230 GLT--DGFAPTLDSFKAIMWDFV-PYFRMKRHRHPIPTQGLAVSRVVGIELMEQKQAV 286  
 DB 209 KLLRFDFDFPLL--FSVLFPFLTPVYEML-----NICMPFKDSIEFFKK----F 252  
 QY 287 LGSASDAQVKKDQVGRDILSLVRA-NIAANLPSQKLSDEVLAAQISNLLFAGYETSS 345  
 DB 253 VDMKESRLDSKQHRVDFLQLMNHNNSKDKVSHKALSDMEITAQSIIFIFAGYETTS 312  
 QY 346 TVLTMPHRLSEDKAVODKLRBEICQI--DTDMPTLDELNALPYLEAFVKESLRDLPSP 403  
 DB 313 STLSTLHSLATHEDFIQKKLQDBIDEALPNKAPPTVDTVMEMEYLDMLVNLTLRYPIAN 372  
 QY 404 YANRECLKDSDFIPLAEPVIGRDSGVINEVRITKGTMMVMLPLFNINRSKFIYGEDAEPR 463  
 DB 373 RLERVCKKQVE-----LNGVYFKGSTVWIPSYALHHDPOHSE-PEEFQ 416  
 QY 464 PERLEWDTVDTLSNS-TEAPYGHQASIFSGPRACFGHRFAVAEMKAPFLVTLRRVQEPIL 522  
 DB 417 PERPSKENKGSIDPVYVLPFG-----NGPRNCLGMRFALNMKLALTIMQNFSPQPC- 469  
 QY 523 SHPEYEHITLIISRPRTVGEEK 544  
 DB 470 ---KETQPLKLSRQGLLOPEK 488  
 RESULT 11  
 C341\_MOUSE  
 ID C341\_MOUSE STANDARD; PRT; 504 AA.  
 AC Q9JMA7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A41 (BC 1.14.14.1).  
 GN CYP3A41.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ddv; TISSUE=Liver;  
 RX MEDLINE=20239668; PubMed=10775455;  
 RA Sakuma T., Takai M., Endo Y., Kuroiwa M., Ohara A., Jarukamjorn K.,  
 RA Honma R., Nemoto N.;  
 RT "A novel female-specific member of the CYP3A gene subfamily in the  
 RT mouse liver.";  
 RL Arch. Biochem. Biophys. 377:153-162(2000).  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum  
 CC (Potential).  
 CC -!- TISSUE SPECIFICITY: Expressed in liver. Also expressed in the  
 CC kidneys of female mice, with traces in the stomach, ovary, and  
 CC heart of female mice and in the testis of male mice.  
 CC -!- DEVELOPMENTAL STAGE: Detected immediately after birth in the  
 CC livers of animals of both sexes, but increased with age in  
 CC females, whereas it was gradually reduced in males, resulting in

```
CC      predominantly female-specific expression in livers.
CC      -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AB033414; BAA95951.1; -.
CC      DR MGI; 1858451; Cyp3a41.
CC      DR InterPro; IPR001128; Cytochrome P450.
CC      DR InterPro; IPR008072; EP450_CYP3A.
CC      DR Pfam; PF00067; P450; 1.
CC      DR PRINTS; PR01689; EP450IICYP3A.
CC      DR PRINTS; PR00385; P450.
CC      DR PROSITE; PS00086; CYTOCHROME P450; 1.
CC      KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC      KW Microsome; Endoplasmic reticulum.
CC      FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC      SEQUENCE 504 AA; 57959 MW; 2BC645B6E9CD48A CRC64;
CC
CC      Query Match      13.7%; Score 391; DB 1; Length 504;
CC      Best Local Similarity 26.2%; Pred. No. 1.8e-15;
CC      Matches 148; Conservative 100; Mismatches 212; Indels 104; Gaps 23;
CC
CC      QY      7 LTGALGLAAPSASIAFFSLY-LAPRRSLYNLQG-PNHT-----NYFTGNFLDIL 55
CC      Db      3 LFSALSLDTWLLAILVLVLYRGTRTHGLFKKQIGPPTPLPFLGTVLNYKGLW---- 58
CC
CC      QY      56 SARTGEHAKYREKYSTL-RFAGIAGAPVLNSTDPKVNHYM-KEAY-----DYPKP 106
CC      Db      59 -----KFDMECYEKYKTNGLFDG--QMPLFVITDPKIMKNLVKCEPSVFTNRREFGPV 111
CC
CC      QY      107 GMAARVLRTAGDVVTAGEAHKRRIMISLSAQAVKSNVPFLEKGMELDKWMD 166
CC      Db      112 GIMSKAISI-----SKDEWKRYALLSTFTSGKLKEMFPPIYQDILVKYLMQE 163
CC
CC      QY      167 AAEKDMVAVGESAGEKATRLTEGVVDVQDWGRATLDVMAIAGFDYKDSLQNKTNELYV 226
CC      Db      164 -AEKG-----KPTMKDVLGAYSIDVITSTSGVWVDSLNNPDPPE 205
CC
CC      QY      227 AFVGL--TGFAFTLSFAKMWDFV-PYFRTWKRHEIPLTQGLAVSRVGIEMBOCK 283
CC      Db      206 KAKGILRVDFDLV--FSVLPFLTPVYEML-----NTCMFPKDSIEFFPK-- 251
CC
CC      QY      284 QAVLGASDQAVKQVQGRDILSLVRANIAANLPESOK-LSDEVLQAQISNLLFAGYE 342
CC      Db      252 --FVNRKESRLSKQKRVDFPLQNMNHNNSKDKSHKALSDMEITAQSIIVIFAGYE 309
CC
CC      QY      343 TSSTVLFWPHRLSEDAKQVDKLRREICQI--DTDMPTLDELNALPYLEAFVKESLRLLDP 400
CC      Db      310 TTSSTLSFTLYCLATHPDIOKKLQEBIDETLPNKAPPTYDTVMEMEYLDVNLNETLVP 369
CC
CC      QY      401 PSPANKECIKDDEFFPLAEPVIGRSGSVINEVRLTKGMVMLPLFNINRSKFIYGEDAE 460
CC      Db      370 IGRLEAFCKDVE-----LNGVTPKGTVMIPSYALHHPDQHPPE-PE 413
CC
CC      QY      461 EFRPERLWDTVDSLNSIEAPYGHQASFGISGRACFGWFAVAEMKAFVTLRVRQEP 520
CC      Db      414 EFQPERFSKNGKSID---PYLYM-PFGIGPRNCIGMRFAPMTMKLALTQVMQNFSPQ 468
CC
CC      QY      521 IISHPEVHEHTLISRPVIGREK 544
CC      Db      469 C-----QETQPLKLSROGLLQPEK 488
CC
CC      RESULT 12
CC      CP37 HUMAN
CC      ID CP37 HUMAN
CC      AC P24462;
```

```
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DE      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Cytochrome P450 3A7 (EC 1.14.14.1) (CYFIIIA7) (P450-HFLA).
GN      CYP3A7.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Fetal liver;
RX      MEDLINE=89255154; PubMed=2722762;
RA      Komori M., Nishio K., Ohi H., Kitada M., Kamataki T.;
RT      "Molecular cloning and sequence analysis of cDNA containing the
RT      entire coding region for human fetal liver cytochrome P-450.";
RL      J. Biochem. 105:161-163(1999).
RN      [2]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=21163842; PubMed=11266076;
RA      Gellner K., Eiselt R., Huestert E., Arnold H., Koch I., Haberl M.,
RA      Deglmann C.J., Burk O., Buntefuss D., Escher S., Bishop C.,
RA      Koebel H.-G., Brinkmann U., Klenk H.-P., Kleine K., Meyer U.A.,
RA      Wojnowski L.;
RT      "Genomic organization of the human CYP3A locus: identification of a
RT      new, inducible CYP3A gene.";
RL      Pharmacogenetics 11:111-121(2001).
CC      -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC      monooxygenases. In liver microsomes, this enzyme is involved in an
CC      NADPH-dependent electron transport pathway. It oxidizes a variety
CC      of structurally unrelated compounds, including steroids, fatty
CC      acids, and xenobiotics.
CC      -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC      oxidized flavoprotein + H(2)O.
CC      -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC      -!- INDUCTION: P450 can be induced to high levels in liver and other
CC      tissues by various foreign compounds, including drugs, pesticides,
CC      and carcinogens.
CC      -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC      -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC      NOTE=CYP3A7 alleles;
CC      WWW="http://www.imm.ki.se/CYPalleles/cyp3a7.htm".
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; D00408; BAA00310.1; -.
CC      DR EMBL; AF280107; RAG32289.1; -.
CC      DR PIR; JX0662; JX0662.
CC      DR HSSP; P14779; LJPZ.
CC      DR Genew; HGNC:2640; CYP3A7.
CC      DR MIM; 605340; -.
CC      DR GO; GO:0015034; F:cytochrome P450 activity; TAS.
CC      DR InterPro; IPR001128; Cytochrome P450.
CC      DR InterPro; IPR008072; EP450_CYP3A.
CC      DR Pfam; PF00067; P450; 1.
CC      DR PRINTS; PR01689; EP450IICYP3A.
CC      DR PRINTS; PR00385; P450.
CC      DR PROSITE; PS00086; CYTOCHROME P450; 1.
CC      KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC      KW Microsome; Endoplasmic reticulum.
CC      FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC      SEQUENCE 503 AA; 57470 MW; 087CCED9BAC314C CRC64;
CC
CC      Query Match      13.5%; Score 385.5; DB 1; Length 503;
CC      Best Local Similarity 29.0%; Pred. No. 3.8e-15;
CC      Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;
```



```
QY 13 LAAPSW--ASTAPSLYL-APRSSLNVLQ-PMHTNY-FTGNFLDILSARTG-----E 61
DQ 7 LAVETLLAVSLILLYGTRHGLFKKIGIPGTPPLPLGN---ALSFRKGYWTFDME 63
QY 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFVNHVM-KEAYD-----YKPPGMAARVL 113
DQ 64 CYKRYKRWG-----YDCQPMALITDPDMIKTLVKECYSYFTNRFPFGVGMQNAI 118
QY 114 RIATGDDGVVTAAGAHKRRIRIMPSLSAQVSMVPIFEKGMELVDKMDAEKDMA 173
DQ 119 SI-----AEDBEWKRIISLSTFTSGKLEMPPIIAQYQDVLVRLRREA----- 164
QY 174 VGESAGEKKAATLET-EGVDVQKDWVGRATLDVWALAGEDYKSDSLQNKTNLYVAVVGLT 232
DQ 165 -----ETGKRVILKLVFGAYSMDVITSFGVSIQSLANPQD-----PPEVNT 207
QY 233 D---GFAPTLDSFKAIMWDFVFPYRTMKRRHEIPLTQGLAVS---RRVGIELMEQKKQAV 286
DQ 208 KKLIRFNP-LDPFVLSIKVF-PFL-----TPILEALNITVFPKRVISFLTKSVKQIK 257
QY 287 LGSASDAQVKKQVQGR---DILSLVVRANITANLPSQKLSDEEVLQISNLLFRQYETS 344
DQ 258 EGRL-----KETQKHVRDQLQMLDSQNSQSETHKALSDLELWQAQSIIFPAGYET 310
QY 345 STVLTMFHRSLSEKAVQDKLREICQIDTDM-----PTLDELNALPYLEAFVKESLRLD 399
DQ 311 SSVLSFIIYELATHPDVQCKVQKE---IDTVLPNKAPPTVDTVLQLEYLDMVNVNLTRLF 367
QY 400 PPSVYANRECLKDEDFIPLAEPVIGROGSVINEVRIKGTWMLPFLFNINRSFVIGEDA 459
DQ 368 PYAMELERVCKDVE-----INGFIPKGVVVMIPSVLVLHDP-KWTTP 411
QY 460 BEFRPERVLEVDVSLNS-IEAPYHQASFTSGPRACFGWRFAVAKFPLFVTLRRVQF 518
DQ 412 EKFLPERFSKKKNKNDIYIITPG-----SGPRNCIGMFEALVNMKLALVRVQLNFSF 465
QY 519 EP 520
DQ 466 KP 467

RESULT 13
CP35 HUMAN
ID CP35 HUMAN STANDARD; PRT; 502 AA.
AC P20815; Q9HB56;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome P450 3A5 (EC 1.4.14.1) (CYP11A5) (P450-PCN3).
GN CYP3A5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=89278095; PubMed=2732228;
RA Aoyama T., Yamano S., Waxman D.J., Iapenson D.P., Meyer U.A.,
RA Fischer V., Tyndale R., Inaba T., Kalcow W., Gelboin H.V.,
RA Gonzalez F.J.
RT "Cytochrome P-450 hPCN3, a novel cytochrome P-450 IITA gene product
RT that is differentially expressed in adult human liver. cDNA and
RT deduced amino acid sequence and distinct specificities of cDNA-
RT expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and
RT cyclosporine."
RL J. Biol. Chem. 264:10388-10395(1989).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
RN [3]
SEQUENCE OF 1-106 FROM N.A.
RP MEDLINE=21163842; PubMed=11266076;
RA Gellner K., Eisel R., Huster E., Arnold H., Koch I., Haberl M.,
RA Deglmann C.J., Burk O., Buntfuss D., Escher S., Bishop C.,
RA Kosse H.-G., Brinkmann U., Klenk H.-P., Kline K., Meyer U.A.,
RA Wojnowski L.;
RT "Genomic organization of the human CYP3A locus: identification of a
RT new, inducible CYP3A gene."
RT Pharmacogenetics 11:111-121(2001).
RN [4]
SEQUENCE OF 1-24 FROM N.A.
RP MEDLINE=96152836; PubMed=8569713;
RA Schuetz J.D., Schuetz E.G., Thottassery J.V., Guzelian P.S., Strom S.,
RA Sun D.;
RT "Identification of a novel dexamethasone responsive enhancer in the
RT human CYP3A5 gene and its activation in human and rat liver cells."
RL Mol. Pharmacol. 49:63-72(1996).
RN [5]
VARIANT CYP3A5*2
RP MEDLINE=96192071; PubMed=8619878;
RA Jounaidi Y., Hyrailles V., Gervot L., Maurel P.;
RT "Detection of CYP3A5 allelic variant: a candidate for the polymorphic
RT expression of the protein?";
RL Biochem. Biophys. Res. Commun. 221:466-470(1996).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP3A5 alleles;
CC WWW="http://www.imm.ki.se/CYPalleles/cyp3a5.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04813; AAA02993.1; --
CC EMBL; BC033862; AAH33862.1; --
CC EMBL; AF280107; AAG32288.1; --
CC EMBL; L35912; AB000083.1; --
CC EMBL; A34101; A34101.
CC HSP; P14779; 1JFZ.
CC Genew; HGNC:2638; CYP3A5.
```

```

DR MIM; 605325; -.
DR GO; GO:0005792; C:mitochondrion; TAS.
DR GO; GO:0015034; F:cytochrome P450 activity; TAS.
DR GO; GO:0004497; F:monooxygenase activity; TAS.
DR GO; GO:0008202; P:steroid metabolism; TAS.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Polymorphism.
FT METAL 441 441 IRON (HEME AXIAL LIGAND).
FT VARIANT 398 398 T -> N (in allele CYP3A5*2).
FT /FTID=VAR_008365.
SQ SEQUENCE 502 AA; 57108 MW; D5A2302E2633E717 CRC64;
Query Match 13.5%; Score 385; DB 1; Length 502;
Best Local Similarity 28.1%; Pred. No. 4.1e-15;
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

DR 13 LAAPSW--ASIAFFSYL-APRSSLYNQ-PNHTNY-FTGNFLDILSARTG-----E 61
DR 7 LAVETWLLAVSLVLLYLYGTRTHGLFKELGIPGTPLELGN---VLSYRQGLWKFDTE 63
DR 62 EHAKYREKYGSTLRFAAGIAGAPVLNSTDKPVNHYM-KEAYDYPKPGMAARVLRIATG-- 118
DR 64 CYKKYKGMGT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRSRLGPV 111
DR 119 ----DGVVTAEGBAHRHRIMIPSLSAQAQVSMVPFLEKGMELVDQMDAAEKMAV 174
DR 112 GFWKSALSLAEDDEWKIRSLSPFTSGKLKEMFPILIAQYGDVLVRLRRE-AEKG--- 167
DR 175 GESAGEKKATRLTEGVYKQVGRATLVMAALGFDYKDSLQ-----KTNELYVAF 228
DR 168 -----KPVTLKIDFGAYSMVDITGTSGVNIIDSLNPPQDPFVESTKXFLK 213
DR 229 VGLTDGFAPLDSFKAIMDVFYFRTMKRHEIPTQGLAVS---RRVGIEMEQKQAV 286.
DR 214 ----GFLDPL-FLSILPFL-----TPVFEALNVSLFPKDTINFLSKS---- 252
DR 287 LGSASDAVDKQVQGRDILSLVRANIANIPESQKLSDEVLQAIQNLIFAGVEISST 346
DR 253 VNRKKSRLNDKQKRLDQLMLDQNSKETESHKALSDELAAQSIIFIFAGVEITSS 312
DR 347 VLTWPHRLSEDKAVODKLEETICQI--DTPMPTDELNALPYLEAFVKESLRDPPSPY 404
DR 313 VLSFTLYELATHPDVQVKLOKEIDAVLPNKAPPTYDAVQMEYLDWVYNEILRFFVAIR 372
DR 405 ANRECKDEDFIPLAEDVIGRDSVINEVRIKGTWMLPLFNIRSKFIYGEDAEFRP 464
DR 373 LERTCKKDVE-----INGVFIPKGMVVIPTVALHDP--KYWTEPEEERF 416
DR 465 ERWLEDVTDLSNS-IEAPYGHQASFTSGPRACGWFPAVAKELFVTLRRVQEP 520
DR 417 ERFSKK-KDSIDPIYTPFG-----TGPRNCIGMFALMNNKALIRVLQNFSEKP 466

RESULT 14
ID CP3S_BOVIN
AC P79102;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A28 (EC 1.14.14.1) (CYP11A28).
GN CYP3A28.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;

```

```

[1]
RN RP
RC STRAIN:Hereford; TISSUE=Liver;
RA Natsuhori M., van Raak M., Lichtenberg M., Kleij L., ten Berge D.,
RA Zweers-Zeilmaker W.M., de Groene E.M., van Miert A.S.J.P.A.,
RA Witkamp R.F., Horbach G.J.M.J.;
RT "Isolation of a bovine full length cytochrome P450 (CYP3A) cDNA
sequence and its functional expression in V79 cells.";
RL Environ. Toxicol. Pharmacol. 3:17-24(1997).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y10214; CAAT1266.1; -.
DR HSSP; P14779; IJFZ.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT SEQUENCE 507 AA; 58152 MW; FDFD8465CC96DF57 CRC64;
Query Match 13.5%; Score 385; DB 1; Length 507;
Best Local Similarity 25.5%; Pred. No. 4.1e-15;
Matches 140; Conservative 94; Mismatches 132; Indels 122; Gaps 22;

QY 38 LQGNHTNYFTGNFLDILSARTGEEHAKYR-----EKYGSTLRFAAGIAGAPVLNS 87
DB 38 IGPFRVPYF-----GSTMAVHKGIPEFDNQCFKKGKMGFYE-GRQPLAI 84
QY 88 TDPKVENHYM-KEAYD-----YKPGMAARVLRIATDGVVTAEGEAHRRHRIMIPS 139
DB 85 TPDPIIKTVLVEKCVSVFTNRIRIFGPMIKYALSLAW-----DEQWKRIETLLSPA 136
QY 140 LSAQAQVKSMPVIFLEKGMELVDQMDAAEKDAAVGSAGEKKAATRLTEGVYKQVWGR 199
DB 137 FTSGKLKEMFPFIIGQYGDMLVRLRKE-AEKG-----NPNVNMKDMFGA 178
QY 200 AILDWALAGDPYKSDSLQNKTNELYAVFGLTGDGAPTLDSF--KAIMMDEV-PYERTM 256
DB 179 YSMDEVITGTAFGNIDSL--NNPHDPFVEHSKNLRLRFP-FDPFILSLILFFLNPVEIL 236
QY 257 KRRHEIPTQGLAVSRVRGIELMEQKQAVLGSASDAQVDKKQVQGRDIISLVANIAA 316
DB 237 -----NITLFPKSVDFTKSVKKI---KESRLTDKQNRVLDLLQLMINSQNSK 282
QY 317 NLPEQSOKLSDEVLQAIQNLIFAGVETSTVLTWPHRLSEDKAVQDKLREIICQITDM 376
DB 283 EIDNHKALSDIELVAQSTIFIFGGYETTSLSFIYELTTHPHVQVKLQEE---IDATP 339
QY 377 -----PTLDELNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAEPVIGRDSVIN 431

```

Db 340 PNRAPTYDALVQMEYLDVMVNETLRFIAGRLSERVCKOVE-----IH 384  
 QY 432 EVRITKGTVMLEPLFNINRSKFTYGDABEERPERWLEDVTDLSNS-IEAPYGHQASFIS 490  
 Db 385 GVPIKGTGTVLPLFLVHNHPPELWPE-PEEPRPERFSKNNKDSINPVVLPFG-----T 437  
 QY 491 GPRACGWRFAVAKAFLEFVTLRRVQPEPI-----ISHREYEHITLISRP 537  
 Db 438 GPRNCLGMRFAIMNIKALVRILONFSPKCKETQIFLKLYTQGLTQPEQPVILKVW--P 495  
 QY 538 RIVGREKE 545  
 Db 496 RGLGPQVE 503

RESULT 15  
 CP3L CALJA STANDARD; PRT; 503 AA.  
 AC O18953;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A21 (EC 1.14.14.1) (CYP11A21) (P450 CM3A-10).  
 GN CYP3A21.  
 OS Callithrix jacchus (Common marmoset).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;  
 CC Callithrix.  
 CC NCBI\_TaxID=94483;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=97223367; PubMed=9056237;  
 RA Igarashi T., Sakuma T., Isogai M., Nagata R., Kamataki T.;  
 RT "Marmoset liver cytochrome P450s: study for expression and molecular  
 RL cloning of their cDNAs."  
 CC Arch. Biochem. Biophys. 339:85-91(1997).  
 CC CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC CC -!- INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR HMBL; D31921; BAA22156.1; -.  
 DR HSSP; P14779; 1JPZ.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00667; P450; 1.  
 DR PRINTS; PR01689; EP450ICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 FT Microsome; Endoplasmic reticulum.  
 KW METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 503 AA; 57564 MW; 9BFD421D72C76D6A CRC64;

Query Match 13.2%; Score 377.5; DB 1; Length 503;  
 Best Local Similarity 27.1%; Pred. No. 1.1e-14;

Matches 148; Conservative 90; Mismatches 184; Indels 125; Gaps 26;  
 QY 13 LAAPSW--ASIAFFSLYADPRRS--SLYNLQG--PHNTNY--FTGNFLDILSARTG-----E 61  
 Db 7 LAVETWLLAVSLVLYLYGTHSHGLFKLGIPOGPTLPFLGT---VLYYRQGFWEKFDME 63  
 QY 62 EHAKYREKYGSLRREAGIAGAPVLNSTDPKVFNVHM--KEAYD-----YPKPGMAARVL 113  
 Db 64 CYKKYKGMWGI---YDG--RQVLAITDPTNIKTVLKECYSVFTNRPRFPVGFVGMKSAI 118  
 QY 114 RIATDGVVTAEGEAHGRHRIMIPSLSAQAVKMWPIFLKGMELVDKMDAABKDMA 173  
 Db 119 STAQDD-----EWKQIRSLSPFTFTSGKLKEMVPIIAQYGEVLVRNLRRE--AEKG-- 167  
 QY 174 VGESAGEKKAETEGVDVKDWGVRATLDVVALAGEDYKSDSLQNKTNELYVAFVGLTD 233  
 Db 168 -----KPINMKDIFGAYSMVDVITGTSFGVNIIDSLNNPD----- 201  
 QY 234 GFAPTLDSFKAIM--WDFV--PYFRVTKRRHEI--PLTQGLAVSR-----RVGIEMLEQ 281  
 Db 202 ---PFVESTKLLRFDLDPDFLPPFLSITIPFPLTILEALNISMFPDSDSTSLRKGSIKRIE 258  
 QY 282 KKOAVLGSASDAQVDKDVQGR--DILSLVYRANITANLPESOKLSDBEVLAAQISNLLFA 339  
 Db 259 SR-----LKOYTHKRVDFLQMLDMSQNKETESDKALSDELVAQSIIFIFA 305  
 QY 340 GYETSSVTILTMFHLRSEDKAVQDKLREEICQIDTDM-----TLDELNALPYLEAFVKE 394  
 Db 306 GYETTSSTLSFIMYELATHPDVQKQLEB---IDAVLPNKAPATYDVLQMEYLDVMVNE 362  
 QY 395 SLRLDPPSYANRECLKDEDFIPLAEPIVGRDGSVINEVRITKGTVMWMLPLFNINRSKFI 454  
 Db 363 TLRLLFLAVRLSERVCKOVE-----INGVFIKGVVVMIPSYALHYDP-K 406  
 QY 455 YGDEADEERPERWLEDVTDLSNS-IEAPYGHQASFISGPRACFGWRFAVAKAFLEFVTL 513  
 Db 407 YWTEPEKFLPERFSKNNKNDIDPIYITPFG-----TGRNCGWRFAVMNKLALIRVL 460  
 QY 514 RRVQFEP 520  
 Db 461 QNFSFKP 467

Search completed: April 2, 2004, 13:59:09  
 Job time : 13 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	419	14.7	593	2	JF6441	2	probable cytochrom
2	406	14.2	503	2	FC4702	2	cytochrome P450 3A
3	404	14.2	503	2	S14275	2	steroid 6beta-mono
4	404	14.2	501	2	A34236	2	cytochrome P450 3A
5	400	14.0	526	2	T02191	2	cytochrome P450 ho
6	399	14.0	504	2	A29410	2	cytochrome P450. g
7	398.5	14.0	503	2	A29815	2	cytochrome P450 3A
8	398	14.0	503	2	S50211	2	cytochrome P450 3A
9	395.5	13.9	503	2	S28168	2	nifedipine oxidase
10	394	13.8	504	2	A60564	2	cytochrome P450 3A
11	395.5	13.5	503	2	JX0062	2	cytochrome P450 3A
12	395	13.5	502	1	A34101	1	cytochrome P450 3A
13	379	13.3	501	2	A29487	2	cytochrome P450 3A
14	370	13.0	504	2	A22631	2	cytochrome P450 3A
15	368	12.9	504	2	S50892	2	cytochrome P450 3A
16	365	12.8	520	2	H84663	2	probable cytochrom
17	363.5	12.7	511	2	T00864	1	cytochrome P450 ho
18	362.5	12.7	523	2	B96562	2	probable cytochrom
19	359.5	12.6	511	2	S66472	2	cytochrome P450 4B
20	357	12.5	502	2	JX0334	2	cytochrome P450 3A
21	355.5	12.5	512	2	A66595	2	hypothetical prote
22	350.5	12.3	1054	1	A69575	1	NADPH-ferrithemopro
23	350	12.3	526	2	CA4533	2	cytochrome P450 4F
24	349.5	12.3	511	1	O4HUB1	1	cytochrome P450 4B
25	349.5	12.3	524	2	S29723	2	cytochrome P450 4F
26	346.5	12.1	511	1	B40164	1	cytochrome P450 4B
27	346.5	12.1	524	2	JC7594	2	cytochrome P450 en
28	346.5	12.1	524	2	JC7598	2	cytochrome P450 en
29	346	12.1	537	2	T02450	2	probable cytochrom

54	Qy	ILSARTG-BEHAKY-REKVGSTLRFAGI--AGAPVLNSTDPKVNHHM-KEAY-----101
50	Db	ILAYRKGFEWFDKYCHHKY---KWLGLYDGRQPPLAITDPDIIKTVLVKECYSTFNER106
102	Qy	DYFKPGWAARVLRIATGDGVTAETGEAHHKRHRIMIPSLSAQAVKSVPIFLEKGMELVD161
107	Db	NFGPVGILKKAISI-----SEDEWKRIPALLSPFTTSGKLKEMFPPI-----149
162	Qy	KMMEDAAEKDMAYGESAGEKKATRLJETEGVDVKWVGGRATLDVNMALAGFYDKSDSLQNT221
150	Db	NQYTDMLVRNRQSGSEKXP-----TSMKDFGAYSMDVITATATSGVNVDSLNNPQ200
222	Qy	NELYVAFVGLTDFGATLDSFKAIM-WD-FVYPYFRTMKRHEI-PLTQGLAVSR-----272
201	Db	D-----PFVKKKLKFDFIEDFLFSLVTLFPFLPLPEALNVSMFPEDVI246
273	Qy	---RVGIELMEQKQAVLGASDAQVKKDVQGRDILSLVRANIATANLPESQKLSDEEV329
247	Db	DFPKTSVERMKENR-----MKEKEKQMDFLQLMINSQNSKVDSHKALSDEVI295
330	Qy	LAQISNLLFAGYETSTVLITWPHRLSEDKAVQDKLRBEICQITDMP-----TLDELNA384
296	Db	VAGSVIPFAGYETTSALSFLVLLIAHPDIQKKLQDE--IDAALPNKAHATYDTLLQ352
385	Qy	LPYLEAFVKESRLDPPSPYANRECLKDSDFPLAEPIVGRDGSVINEVRTIKCTWMVLP444
353	Db	MEYLDVMVNETLLYPIAGRLERVKCTDVE-----INGVIFPKGTVMVNIIP397
445	Qy	LFINNRSKFTYGDCAEEFPERLWEDVTDLSLNSIEAPYCHQASTISGPACFCGWRFFAAE504

cytochrome P450 3A6 (version 2) - rabbit  
 N:Alternate names: cytochrome P450 3C  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 28-Jul-2000  
 C:Accession: A34236  
 R:Portenza, C.L.; Pendurthi, U.R.; Strom, D.K.; Tukey, R.H.; Griffin, K.J.; Schwab, G.E.;  
 J. Biol. Chem. 264, 16222-16228, 1989  
 A:Title: Regulation of the rabbit cytochrome P-450 3C gene. Age-dependent expression and  
 A:Reference number: A34236; MUID:89380226; PMID:2777787  
 A:Accession: A34236  
 A:Molecule type: mRNA  
 A:Residues: 1-501 <POT>  
 A:Cross-references: GB:J05034; NID:g164829; PIDN:AAA31178.1; PID:g164830  
 C:Genetics:  
 A:Gene: CYP3A6  
 C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;  
 F:300-462/Domain: cytochrome P450 homology <P45>  
 F:440/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.1%; Score 402; DB 2; Length 501;  
 Best Local Similarity 27.8%; Pred. No. 1.2e-18;  
 Matches 156; Conservative 88; Mismatches 200; Indels 118; Gaps 24;

QY 19 ASIAPFSLYLAPRRSSLYNQGPNHTNY-FTGNFLDILSARTG--BEHAKYREKYGSTLR 75  
 DB 15 ASLVLLYLYGSTHGLFKKMGIPGPTPLPFIGT---ILEYRKGIWDPIECRKYG--K 68

QY 76 FAGI--AGAPVNLSTDPKVFNVHM-KEAY-----DYPKPGMAARVLRATGCGVVTAE 125  
 DB 69 MWGLFDGQPLVITDPDMIKTLVLCYCVSYFTNRRSGFVGMFKKAVSI-----SE 120

QY 126 GEAKHRRIRIIPSLSAQAQSMVPIFELEKGMELVDKMDAEAKMAVGESAGEKKATR 185  
 DB 121 DEDWKRVTLSPTFTSKLKEMLPIIAQYGDVLV-KNLQAEAKG-----165

QY 186 LETEGVDYKDWGRATLDVMAAGFDYKSDSLQNKTN-----LYAVFGLTDFGAP 237  
 DB 166 ---KEVDLKEIFGAYSMVDITGTSFGVNIDSLRNPQDPFVKVNRLLKFSF-----FDP 216

QY 238 TLDSEKAIWMDVFPVFRTRMKRRIEPLTQGLAVSRVYGIEMQKQAVLSASDAQVK 297  
 DB 217 LLSI-----TLFPFL-----TPFEALHIS-MPFKVMDFLTSVEKIKDDRLKK 262

QY 298 KDVGQRDILSLVRANAIANLPSQKLSDEEVLQAISNLLFAGYETSTVLTWMFRLSE 357  
 DB 263 QK-RRVDFLQLMINSQNSKEIDSHKALDDIEVVAQSIILFAGYETSTSLSPIMHLLAT 321

QY 358 DRVQDKLREECQIDTDP-----TLDELNALPYLEAFYKESLRDPPFPYANRECLKD 412  
 DB 322 HPDVQOKLQEE---IDTLNPKELATYDVLVKNYDLMVNETURLYPIAGRLERVCCKD 378

QY 413 EDFIPLAEFVIGRDSVINEVRITKGTVMPLFNINRSFYGEDAEFFRPERLWEDVT 472  
 DB 379 VD-----INGFIPKGTVMPTVALHRDPQHWE-PDEFERPERFSKKN 422

QY 473 DSLNS-IEAPYGHQASFTSGPRACGWPFAVAKMAFLVTLRVQPE-----PI---521  
 DB 423 DNINPYIYHPFG-----AGPRNCLGMRFALMNKIALVRLMQNFSPLCKETQVPLKLG 476

QY 522 ---ISHPEVEHETLIIISPRIV 540  
 DB 477 KQGLLOQPEKPIVKKVSRDGI 498

RESULT 5  
 T02191  
 cytochrome P450 homolog F14M4.21 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
 C:Accession: T02191; D84909  
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rott

submitted to the EMBL Data Library, September 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.  
 A:Reference number: Z14609  
 A:Accession: T02191  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-526 <RCU>  
 A:Cross-references: EMBL:AC004411; NID:g3522932; PID:g3522945  
 A:Experimental source: Cultivar Columbia  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84909  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-526 <STC>  
 A:Cross-references: GB:AE002093; NID:g3522945; PIDN:AAC34227.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g46960; F14M4.21  
 A:Map position: 2  
 A:Introns: 93/1; 168/3; 254/2; 375/3  
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:322-493/Domain: cytochrome P450 homology <P45>  
 F:471/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 400; DB 2; Length 526;  
 Best Local Similarity 23.6%; Pred. No. 1.7e-18;  
 Matches 140; Conservative 109; Mismatches 231; Indels 114; Gaps 21;

QY 1 MFIVLLTGTALGAARSWASIAFSLYLAPRRSSLYNQGPNHTNYFTGNF-----51  
 DB 9 LIVLILI---IGURIFKAFMLVWHPVLRLLKNOGIGSPNY-RIFYGNLSEIKOMKE 64

QY 52 --DLISARTGE-----EHAKYREKYSTLRFAGIAGAPVNLSTDPKVFNVHWEAYD 102  
 DB 65 SHLSILDPSSNDIPPRILPHYQKMSQYGETFLYWN-GTEPRICISDPELAKTMSNKLIG 123

QY 103 -YPPGMAARVLRATGCGVVTAEAGHAKHRRIMIPSLSAQAQSMVPIFELEKGMELVD 161  
 DB 124 FVTSKARPAEAKLVGSKGLVFEIGADWVRHRLINPAFISIDELKIMTVMVDCIL----179

QY 162 KQMDAAEKDMAVGESAGEKATRLTEGVVDKDWVG-----RATLDVMAAGFDYKSDSL 217  
 DB 180 KWLSE-----WRKESTKEETHPKIKKENNEEFORLTADIIATSAFG-----221

QY 218 QNKTNELVAVFGLTDFGAPFLDSEKAIWMDVFPVFRTRMKRRIEPLTQGLAVSRVIGTE 277  
 DB 222 -----SSYVEGIEVFRSQM-ELKRCYTTSLNQVSIPTQYLPFSNIRW 265

QY 278 LMEQKQAVLSASDAQVDKQVGRDILSLVRANAIANLPSQKLSDEEVLQAISNLL 337  
 DB 266 KLERKMDNSIKRIISSRLQSKSDYGDLLGILLKAYNTEG--KERKMSIEEIIHECRTEF 323

QY 338 FAGYETSTVLTWMFRLSEDKAVQDKLREEI---C-QIDTMDPTLDLNLALP-----YL 388  
 DB 324 FGGHETSNLAWTMTLSLHQDKQKREIEFPKCGKTKPDSETFSKLKUKPVVFLQ 383

QY 389 EAFVKESLRDPPFPYANRECLKDEDFIPLAEFVIGRDSV---INEVRITKGTVMVLEL 445  
 DB 384 NMVIMESIRLYGP-----VSALAREASVNIKGLDLBPKGTTVWIFL 425

QY 446 FNINRSKFIYGEDAEFFRPERLWEDVTDSLSNSTEAPYGHQASFIQSPRACFGHFRFAEM 505  
 DB 426 LKMHSDKTLWGSADKFNPRFANGVSRANHPNA-----LLAFSVGPRACIQGNFWIEA 481

QY 506 KAPLVTLRRVQPEFIISHPEYEH---ITLIIISPRIVGREKEGYQMRLOVKVE 557  
 DB 482 KTVLTMLQRRFISLQCD--EYKHTPVDNVTIQFQ-----YGLPVMVLPLE 525

RESULT 6

A29410  
cytochrome P450, glucocorticoid-inducible, hepatic - human  
N:Alternate names: cytochrome P450 3A3; cytochrome P450 H1p  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1989 #sequence revision 31-Mar-1989 #text\_change 28-Jul-2000  
C:Accession: A29410; A25170; A25450  
R:Molowa, D.T.; Schuetz, E.G.; Wrighton, S.A.; Watkins, P.B.; Mendez-Picon, R.; Watkins, P.B.; Wrighton, S.A.; Maurel, P.; Schuetz, E.G.; Mendez-Picon, G.  
Proc. Natl. Acad. Sci. U.S.A. 83, 6310-6314, 1986  
A:Title: Identification of an inducible form of cytochrome P-450 in human liver.  
A:Reference number: A25170; MUID:85298342; PMID:3898085  
A:Accession: A25170  
A:Molecule type: protein  
A:Residues: 2-21 <WAT>  
A:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; heme; iron; liver; metalloprotein; monooxy  
F:303-465/Domain: cytochrome P450 homology <P45>  
F:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 399; DB 2; Length 504;  
Best Local Similarity 28.4%; Pred. No. 1.9e-18;  
Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;  
QY 13 LAAPSW--ASIAFFSLYLAPRS-SLYNLQG-PNHTNY-FTGNFLDILSARTG-----E 61  
DB 7 LAMETLLLAVALSVLLYLYGTHSHGLFKLGGPGTTPPLFLGN---ILSYHKGCFMDE 63  
QY 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYD-----YKPGMAARVL 113  
DB 64 CHKKYKQWGF---YDQ--QQPVLAITDPDMIKLVKCEYSVFTNRRPFGVGMKSAI 118  
QY 114 RIATGDGVTAEGEAHRRHRIIMPSSAQAVKSVVPIFLEKGMELVDKMWEDAAEKDMA 173  
DB 119 SI-----AEDEEWKRLSLSPFTSGKLEKWPPIIAQGDVVLVNLRR----- 163  
QY 174 VGSAGEKATRLTEGVDVQWVGVRATLDVVALAGFDYKSDSLQNKTNELYVAVGLTD 233  
DB 164 -----RETGKPVTLKQVFGAYSMDVITSSGFGVNVDSLNNPD----- 201  
QY 234 GPAPTLDLFAIM-WDFV-PYFTMK-RRHEIPLTQGL-----AVSRVVG 275  
DB 202 ---PLVENTKKLRFDFLDFPLSITVTFPLIPILVLELVNICVFPREVTNFKAVKMK 258  
QY 276 IELMEQKQAVLGASDAQVKKDQVGRDILSLVRAIANLPESQK-LGDEEVLQIS 334  
DB 259 SRLDGTQHRV-----DFLQMLDHSKNSKETSSHKALSLELVAQSI 301  
QY 335 NLLFAGVETSTVLTMMFHLSDKAVQKLRKEIQI---DIDMTLDELNALPYLEAFV 392  
DB 302 IFIFAGVETTSVSLFIMYELATHPDVQKQLOEIDAVLPNKAPPYDTVLQMEYLDWV 361  
QY 393 KESRLDPPSPYANRECLDKEDFIPLAEPVIGRDSVINEVRITKTMVWMLPLFNINRSK 452  
DB 362 NETLRLFPANRLERVCKDVE-----INGMFIKGVWVMIYSYALHRDP 406  
QY 453 FYIGEDAEFEPERLWEDVTSLSNS-IEAPYHQACFSIGPRACFQWRFAVEMKAFLEV 511  
DB 407 -KYWTEBEKFLPERFSKKNKNDIPYITPFG-----SGPRNCIWMRFALMKNKLALIR 459  
QY 512 TLRRVQREP 520  
DB 460 VLQNFSPKP 468

RESULT 7

A29815  
cytochrome P450 3A4 nifedipine oxidase (EC 1.14.14.-) - human  
N:Alternate names: cytochrome P450 (PCN1); cytochrome P450-HM1  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1989 #sequence revision 30-Sep-1989 #text\_change 28-Jul-2000  
C:Accession: A29815; A15900; A25517; A32199; PX0012; S03851  
R:Gonzalez, F.J.; Schmid, B.J.; Umeno, M.; McBride, O.W.; Hardwick, J.P.; Meyer, U.A.;  
DNA 7, 79-86, 1988  
A:Title: Human P450PCN1: sequence, chromosome localization, and direct evidence through  
A:Reference number: A29815; MUID:88195781; PMID:3267210  
A:Accession: A29815  
A:Molecule type: mRNA  
A:Residues: 1-503 <GON>  
A:Cross-references: GB:M18907; NID:G181373; PIDN:AAA35745.1; PID:G181374  
R:Spurr, N.K.; Gough, A.C.; Stevenson, K.; Wolf, C.R.  
Hum. Genet. 81, 171-174, 1989  
A:Title: The human cytochrome P450 CYP3 locus: assignment to chromosome 7q22-qter.  
A:Reference number: S16900; MUID:89108438; PMID:2563251  
A:Accession: S16900  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-391, 'W', 393-503 <SP2>  
A:Cross-references: EMBL:X12387; NID:G35910; PIDN:CAA30944.1; PID:G35911  
R:Beaune, P.H.; Umbenhauer, D.R.; Bork, R.W.; Lloyd, R.S.; Guengerich, F.P.  
Proc. Natl. Acad. Sci. U.S.A. 83, 8064-8068, 1986  
A:Title: Isolation and sequence determination of a cDNA clone related to human cytochrome  
A:Reference number: A25517; MUID:87041402; PMID:3464943  
A:Accession: A25517  
A:Molecule type: mRNA  
A:Residues: 1-391, 'W', 393-503 <BEA>  
R:Bork, R.W.; Muto, T.; Beaune, P.H.; Srivastava, P.K.; Lloyd, R.S.; Guengerich, F.P.  
J. Biol. Chem. 264, 910-919, 1989  
A:Title: Characterization of mRNA species related to human liver cytochrome P-450 nifed  
A:Reference number: A32199; MUID:89093163; PMID:2463251  
A:Accession: A32199  
A:Molecule type: protein  
A:Residues: 'X', 2-9, 'XXX', 13-15 <BOR>  
R:Komori, M.; Hashizume, T.; Ohi, H.; Miura, T.; Kitada, M.; Nagashima, K.; Kamataki, T.  
J. Biochem. 104, 912-916, 1988  
A:Title: Cytochrome P-450 in human liver microsomes: high-performance liquid chromatogr.  
A:Reference number: PX0012; MUID:89214010; PMID:3243766  
A:Accession: PX0012  
A:Molecule type: protein  
A:Residues: 'X', 2-11, 'X', 13-25 <KOM>  
A:Experimental source: liver microsome  
C:Genetics:  
A:Gene: GDB:CYP3A4  
A:Cross-references: GDB:118782  
A:Map position: 7q22.1-7q22.1  
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metall  
F:302-464/Domain: cytochrome P450 homology <P45>  
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 398.5; DB 2; Length 503;  
Best Local Similarity 28.7%; Pred. No. 2e-18;  
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
QY 13 LAAPSW--ASIAFFSLYLAPRS-SLYNLQG-PNHTNY-FTGNFLDILSARTG-----E 61  
DB 7 LAMETLLLAVALSVLLYLYGTHSHGLFKLGGPGTTPPLFLGN---ILSYHKGCFMDE 63  
QY 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYD---PKPGMAARVLRIAT 117  
DB 64 CHKKYKQWGF---YDQ--QQPVLAITDPDMIKLVKCEYSVFTNRRPFGVGMKSA- 117  
QY 118 GDGVTAEGEAHRRHRIIMPSSAQAVKSVVPIFLEKGMELVDKMWEDAAEKDMAVGES 177  
DB 118 ---ISIAEDEEWKRLSLSPFTSGKLEKWPPIIAQGDVVLVNLRR----- 164

303 IFAGYETTSSALSFALYLLATHPDVQKKLODE--IDAALPNKAPATYDTLLQMEYLDNV 359

392 VKESURLDPSPSYANRECLKDDFTPLAEPVIGRDSGVINEVRITKGMVMLPLENIRS 451

360 VNETHLYPIAGRLERVCKTDE-----INGLFIKGVIVMLTPTFALHKD 404

452 KFTYGSDAEFEPERLEWVDVDTLSNLSIAPYGHQASFISGPRACFGWRFAVAMKAFJLV 511

405 P-KYWPPEPEFPERFKSKNQDSIN---PYMY-LPFGSGPRNCIGMRFALLNMKVALVR 458

512 TLRVQFEPI-----ISHPEYEHITLISPRIVGRE 543

459 VLQNFVTQPKETETPLKLSKGLQLQEPENPLLLKVVSRDETVSDE 503

RESULT 9

S28168

nifedipine oxidase (EC 1.14.14.-) cytochrome P450 3A8 - crab-eating macaque

N/Alternate names: cytochrome P450 CMLC; cytochrome P450 Mknf2; cytochrome P450-MK2

C/Species: Macaca fascicularis (Crab-eating macaque)

C/Date: 28-May-1993 #sequence revision 28-May-1993 #text\_change 28-Jul-2000

C/Accession: S28168; S04509; S36875

E/Komori, M.; Kikuchi, O.; Sakuma, T.; Funaki, J.; Kitada, M.; Kamataki, T.

Biochim. Biophys. Acta 1171, 141-146, 1992

A/Title: Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity of the pri

A/Reference number: S28166; MUID:93129612; PMID:1282830

A/Accession: S28168

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-503 <KOM>

A/Cross-references: GB:S53047; NID:Q263692; PIDN:AAB24952.1; PID:Q263693

R/Ohta, K.; Kitada, M.; Hashizume, T.; Komori, M.; Ohi, H.; Kamataki, T.

Biochim. Biophys. Acta 996, 142-145, 1989

A/Title: Purification of cytochrome P-450 from polychlorinated biphenyl-treated crab-eat

A/Reference number: S04509; MUID:89287352; PMID:2500151

A/Accession: S04509

A/Molecule type: protein

A/Residues: 1-22 <OHT>

R/Ohmori, S.; Horie, T.; Guengerich, F.P.; Kiuchi, M.; Kitada, M.

Arch. Biochem. Biophys. 305, 405-413, 1993

A/Title: Purification and characterization of two forms of hepatic microsomal cytochrome

A/Reference number: S36874; MUID:93384294; PMID:8373178

A/Accession: S36875

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-13 <OHM>

C/Genetics:

A/Gene: CYP3A8

C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology

C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo

E/302-464/Domain: cytochrome P450 homology <P45>

F/442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.9%; Score 395.5; DB 2; Length 503;

Best Local Similarity 28.6%; Pred. No. 3.2e-18;

Matches 155; Conservative 85; Mismatches 187; Indels 115; Gaps 26;

QY 13 LAAPSW--ASIAFFSYLAPRRS-SLYNLQG-PNHINY-FTGNFLDILSARTG-----E 61

DB 7 LAVETWLLAVTLVLLVLYGTHSHGLFKKLIGPGTLPPLGN--ILSVKKGFWTFDME 63

QY 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKFNHVM-KEAYD-----YPKGGMARVL 113

DB 64 CYKKYKVGWGF---YDG--RQPVLAITDPENMIKTVLVKCYSVFTNRRPGPGVFWKNAI 118

QY 114 RIATGDGVVTAEGEAHKKRRIRI:PSLSAQAVKSWVPI:FEKGMELVDKMDA:AEKDMA 173

DB 119 SI-----AEDEWKRISLSPTFTSFKLKEMVPIAKYGDVLRNLRREA----- 164

QY 174 VGESAGEKKATRLT-EGVDVKQDWGRATLDVVALAGFDYKSDSLQNKTNELVAFVGLT 232

DB 165 -----ETGKPVTLKQVFGAYSMDVITSTSGVNI:DSLNNPD----- 201





C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
F:302-464/Domain: cytochrome P450 homology <P45>  
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.5%; Score 385.5; DB 2; Length 503;  
Best Local Similarity 29.08; Pred. No. 1.4e-17;  
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61  
DB 7 LAVETWLLAVSLVLLYGRTHGLFKLIGIPGTPLPFLGN---ALSFRGYWTFDME 63  
QY 62 EHAQYREKYGSTLRPAGIAGAPVLNSTDPKFNHVM-KEAYD-----YPKPGMAARVL 113  
DB 64 CYKIRKVGW-----YDCCQPMALITDPNMTKVLVKECYSVFTNRRPFGPVGFWKNAI 118  
QY 114 RIATGDDGVVTAAGBAHRRIRIMIPSLSAQAQVSMVPIFLEKGMELVDKMDAAEKDVA 173  
DB 119 SI-----AEDEWKIRSLSPFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164  
QY 174 VESAGEKKATRLT-EGVDVKDWVGRATLDVMAAGFDYKSDSLONKTNELVAVPGLT 232  
DB 165 -----ETGKPVTLKHVFGAYSDMTITSTSGVSDSLNPDQ-----PPVENT 207  
QY 233 D---GFAPTLDSFKAIMWDFVFPYRTMKRRHEIPLTQGLAVS---RRVGIELMEQKQAV 286  
DB 208 KKLRFNP-LDPFVLSIKVP-PFL-----TPILEALNITVPPRKVISFLTSVKQIK 257  
QY 287 LGSASDAQVKKVQGR--DILSLVRANIANLPSPQKLSDEVLQAQISNLLFAGYETS 344  
DB 258 EGRLL-----KETQKRVDFLQMLDSQNSKDSETHKALSDELAQAQSIIFIFAGYETT 310  
QY 345 STVLTWFMHRLSEDAVQDKLREICQIDTDM-----PTLDELNALPYLEAFVKSRLRD 399  
DB 311 SSVLSFIYELATPDVQKQVKE---IDTVLPNKAPPTVDTVLQLEYLDMVNETLRLP 367  
QY 400 PPSPYANRECLKDEDFIPLAEPVIGRQGSVINEVRITKGMVWMLPLFNINRSKFYGEDA 459  
DB 368 PVAMRLERVCKKQVE-----INGMFIPKGVVWVIMPSVYLHNDP-KYWTETP 411  
QY 460 EEPFRWLEDVTSLSN-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRLRVQF 518  
DB 412 EKFLPERFSKKKNIDPIYITPFG-----SGPRNCIGMRFALVNKKLALVRLVQNFSF 465  
QY 519 BP 520  
DB 466 KP 467

RESULT 12  
A34101  
cytochrome P450 3A5 - human  
N:Alternate names: cytochrome P450 HLP2  
N:Contains: oxidoreductase (EC 1.1.1.1)  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: A34101; S06491; I52302  
R:Aoyama, T.; Yamano, S.; Waxman, D.J.; Lapenson, D.P.; Meyer, U.A.; Fischer, V.; Tyndal  
J. Biol. Chem. 264, 10388-10395, 1989  
A:Title: Cytochrome P-450 hPCN3, a novel cytochrome P-450 gene product that is diff  
DNA-expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and cyclosporine.  
A:Reference number: A34101; MUID:89278095; PMID:2732228  
A:Molecule type: mRNA  
A:Residues: 1-502 <AOY>  
A:Cross-references: GB:J04813; NID:G181345; PIDN:AAA02993.1; PID:G181346  
R:Schuetz, J.D.; Molowa, D.T.; Guzelian, P.S.  
Arch. Biochem. Biophys. 274, 355-365, 1989  
A:Title: Characterization of a cDNA encoding a new member of the glucocorticoid-responsi  
A:Reference number: S06491; MUID:190025114; PMID:2802615  
A:Accession: S06491  
A:Status: not compared with conceptual translation

A:Molecule type: mRNA  
A:Residues: 1-304, 'P', 306-317, 'F', 319-323, 'D', 325-376, 'G', 378-502 <SCH>  
R:Jounaidi, Y.; Guzelian, P.S.; Maurel, P.; Villarem, M.J.  
Biochem. Biophys. Res. Commun. 205, 1741-1747, 1994  
A:Title: Sequence of the 5'-flanking region of CYP3A5: comparative analysis with CYP3A4  
A:Reference number: I52302; MUID:95110318; PMID:7811260  
A:Accession: I52302  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-24 <RES>  
A:Cross-references: GB:S74699; NID:G786472; PIDN:AA014157.1; PID:G4261857  
C:Genetics:  
A:Gene: GDB: CYP3A5  
A:Cross-references: GDB:118783  
A:Map position: 7q22.1-7q22.1  
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;  
F:302-463/Domain: cytochrome P450 homology <P45>  
F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.5%; Score 385; DB 1; Length 502;  
Best Local Similarity 28.1%; Pred. No. 1.5e-17;  
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61  
DB 7 LAVETWLLAVSLVLLYGRTHGLFKLIGIPGTPLPFLGN---VLSYRQGLWKFDTE 63  
QY 62 EHAQYREKYGSTLRPAGIAGAPVLNSTDPKFNHVM-KEAYDYPKPGMAARVLRATG-- 118  
DB 64 CYKIRKVGWGT---YEG--QLPVLAITDPVIRTVLVECY-----SVFTNRSRLGPV 111  
QY 119 ----DGVTVAAGBAHRRIRIMIPSLSAQAQVSMVPIFLEKGMELVDKMDAAEKDVA 174  
DB 112 GFMKSAISLAEDBEWKIRSLSPFTSGKLKEMVPIIAQYGDVLRNLRRE-ASKG-- 167  
QY 175 GESAGEKKATRLTETEGVDVKDWVGRATLDVMAAGFDYKSDSLON-----KTNELVAVF 228  
DB 168 -----KPTVLKDIIFGAYSDMTITSTSGVSDSLNPDQPFVETSKFLKF 213  
QY 229 VGLTGDGFAPTLDSFKAIMWDFVFPYRTMKRRHEIPLTQGLAVS---RRVGIELMEQKQAV 286  
DB 214 -----GFLDPL-FLSILPPFL-----TPVEALNVSLFPKDTINFLSKS---- 252  
QY 287 LGSASDAQVKKVQGRDILSLVRANIANLPSPQKLSDEVLQAQISNLLFAGYETS 346  
DB 253 VNRMKSRNDKQKRLDFLQMLDSQNSKDETHKALSDELAQAQSIIFIFAGYETS 312  
QY 347 VLTWFMHRLSEDAVQDKLREICQI---DTDMPTLDELNALPYLEAFVKSRLRD 404  
DB 313 VLSFTLYELATHPDVQKQVKEIDAVLPNKAPPTVDVAVVQMEYLDVWVNETLRLFPVAIR 372  
QY 405 ANRECLKDEDFIPLAEPVIGRQGSVINEVRITKGMVWMLPLFNINRSKFYGEDA 464  
DB 373 LERTCKKQVE-----INGVFIKGSMMVVIPTVALHNDP-KYWTETP 416  
QY 465 ERLWLEDVTSLSN-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRLRVQF 520  
DB 417 ERFSKK-KDSIDPIYITPFG-----TGRNCIGMRFALVNKKLALVRLVQNFSF 466

RESULT 13  
A29487  
cytochrome P450 3A6 (version 1) - rabbit  
N:Alternate names: cytochrome P450 3C  
N:Contains: oxidoreductase (EC 1.1.1.1)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text\_change 28-Jul-2000  
C:Accession: A29487  
R:Dalet, C.; Clair, P.; Daujat, M.; Fort, P.; Blanchard, J.M.; Maurel, P.  
DNA 7, 39-46, 1988  
A:Title: Complete sequence of cytochrome P450 3c cDNA and presence of two mRNA species w  
A:Reference number: A29487; MUID:88166352; PMID:3349903

A;Accession: A29487  
A;Molecule type: mRNA  
A;Residues: 1-501 <DAL>  
A;Cross-references: GB:M19139; NID:G165573; PIDN:AAA1430.1; PID:G165574  
C;Genetics:  
A;Gene: CYP3A6  
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase  
F;301-462/Domain: cytochrome P450 homology <P45>  
F;440/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.3%; Score 379; DB 2; Length 501;  
Best Local Similarity 26.2%; Pred. No. 3.8e-17;  
Matches 147; Conservative 91; Mismatches 208; Indels 114; Gaps 22;  
Matches 147; Conservative 91; Mismatches 208; Indels 114; Gaps 22;

Qy 19 ASTAFSLYLAAPRSSLYNLOGNHTNY-FTGNFLDILSARTGEHAKYREKYGTLPRA 77  
Db 15 ASLVLLYLYGTSTHGLFKXGIPGPPPLPFIPTILLEYRKYVINWDFIECRKXYG---KMW 71  
Qy 78 GI--AGAPVLNSTDPKVFNVHM-KEAY-----DYPKPGMARVLRATGDGVVTAEGE 127  
Db 72 GLFDGRQPLMVIITDPMIKTVLKECYSVFTNRRSGPGVGFMKKAVSI-----SEDE 123  
Qy 128 AHKXRRIMTSPSAQAQSMVDPIFEKGMELVDKMDAEDAKOMAVGESAGEKATRL 187  
Db 124 DWKRVRTLLSPTTSKGLKEMPIIAQYGDVLVKNLQBAEASPSST----- 170  
Qy 188 TEGVDVVDKMGVGRATLDVMAAGFDYKSDSLQKNTNE-----LYVAFVGLTDGFAPTL 239  
Db 171 -----LKEIFGASMDVITGTSFGVNIDSLRPQDPFVKVNRLLKPSF-----FDPLL 219  
Qy 240 DSKALMDFVPYFRTMKRRHRIPLTQGLAVRRVGIEMEQKQAVLGSADQAVKDD 299  
Db 220 LSI-----TLFPFL-----TPIEALHIS-MFPKQVMDPLKTSVEKIKDKRQK 265  
Qy 300 VQGRDILSLAVRANIAANLPESQKLSDEVLQISNLPAGVETSGTSLTWMHRLSEDK 359  
Db 266 -RRVDFIQLMINSQNSKEIDSHKALDDIEVVAQSIILPAGYETTSLTSLFIMHLLATHP 324  
Qy 360 AVQDKLREEICQIDTMP-----TLDELNALPYLEAFVKESLRDPPSPYANRECLKDD 414  
Db 325 DVQOKLQEE---IDTLPLNELATYDTLVKMEYLDVNVNETURLYPIAGRLERVKCKDD 381  
Qy 415 FIPLAEPIVIGDSVINEVRITGTMVPLFNINRSKFIYGEDAEERPERNLEDDVTS 474  
Db 382 -----INGTIPKGTIVM-PTVALHRDPQHWTPE-PDEPRPERFSKQNDN 424  
Qy 475 LNS-IEAPYGHQAFISGPRACFGWFAVAKKAFLEFVTLRRVQPE-----PT----- 521  
Db 425 INPIYHPFG-----AGPENCILGMRFALMNIKALVRLMQNFSFKLCKETQVPLKLGQ 478  
Qy 522 -ISHPEYEHITLISPRIV 540  
Db 479 GLLOPEKPIVLKVVSSGII 498

RESULT 14  
A22631  
Cytochrome P450 3A1, pregnenolone 16-alpha-carbonitrile-inducible - rat  
N;Alternate names: testosterone 6beta-hydroxylase  
N;Contains: unspecific monooxygenase (EC 1.14.14.1) cytochrome P450 PCN1  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 29-Aug-1987 #sequence revision 29-Aug-1987 #text change 28-Jul-2000  
C;Accession: A22631; PX0035; S21697; S36137; S27107; S30378; I59218  
R;Gonzalez, F.J.; Nebert, D.W.; Hardwick, J.P.; Kasper, C.B.  
J. Biol. Chem. 260, 7435-7441, 1985  
A;Title: Complete cDNA and protein sequence of a pregnenolone 16-alpha-carbonitrile-indu  
A;Accession: A22631; MUID:85207783; PMID:3638989  
A;Molecule type: mRNA  
A;Residues: 1-504 <GN>  
A;Cross-references: GB:M10161; NID:G203777; PIDN:AAA1035.1; PID:G203778  
R;Negata, K.; Gonzalez, F.J.; Yamazoe, Y.; Kato, R.

J. Biochem. 107, 718-725, 1990  
A;Title: Purification and characterization of four catalytically active testosterone 6b  
nally related forms.  
A;Reference number: PX0032; MUID:90375438; PMID:2398039  
A;Accession: PX0035  
A;Molecule type: protein  
A;Residues: 1-26 <NAG>  
A;Experimental source: liver, Sprague-Dawley male rat  
R;Lechner, M.C.  
submitted to the EMBL Data Library, December 1991  
A;Reference number: S21697  
A;Accession: S21697  
A;Molecule type: mRNA  
A;Residues: 1-206, 'A', 208-212, 'I', 214-231, 'V', 233-504 <LEC>  
A;Cross-references: EMBL:X64401; NID:G56038; PIDN:CAA45743.1; PID:G56039  
R;Ribeiro, V.; Lechner, M.C.  
Arch. Biochem. Biophys. 293, 147-152, 1992  
A;Title: Cloning and characterization of a novel CYP3A1 allelic variant: Analysis of CY  
A;Reference number: S36137; MUID:92117688; PMID:1731631  
A;Accession: S36137  
A;Molecule type: mRNA  
A;Residues: 205-206, 'A', 208-212, 'I', 214-231, 'V', 233-234 <RIB>  
A;Cross-references: EMBL:X64401  
R;Telhada, M.B.; Pereira, T.M.; Lechner, M.C.  
Arch. Biochem. Biophys. 298, 715-725, 1992  
A;Title: Effect of dexamethasone and phenobarbital on run-on transcription rate and CYP  
A;Reference number: S27107; MUID:93037516; PMID:1417000  
A;Accession: S27107  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-24 <TEL>  
A;Cross-references: EMBL:X62086  
R;Cooper, K.O.; Reik, L.M.; Jayyosi, Z.; Bandiera, S.; Kelley, M.; Ryan, D.E.; Daniel,  
Arch. Biochem. Biophys. 301, 345-354, 1993  
A;Title: Regulation of two members of the steroid-inducible cytochrome P450 subfamily (  
A;Reference number: S30378; MUID:93213168; PMID:7681660  
A;Accession: S30378  
A;Molecule type: protein  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-25 <COO>  
Proc. Natl. Acad. Sci. U.S.A. 89, 2145-2149, 1992  
A;Title: Paradoxical transcriptional activation of rat liver cytochrome P-450 3A1 by de  
to primary monolayer cultures of adult rat hepatocytes.  
A;Reference number: I59218; MUID:92196074; PMID:1372436  
A;Accession: I59218  
A;Status: preliminary  
A;Molecule type: protein  
A;Cross-references: GB:M86850; NID:G205919; PIDN:AAA41780.1; PID:G205920  
C;Genetics:  
A;Gene: CYP3A1; P450p  
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase  
F;303-465/Domain: cytochrome P450 homology <P45>  
F;443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.0%; Score 370; DB 2; Length 504;  
Best Local Similarity 26.2%; Pred. No. 1.5e-16;  
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;  
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

Qy 7 LTGALGLAFAFWASTAFSLY-LAPRRSSLYNLOG-----PHTNYFTGNF-LDI 54  
Db 3 LLSALTLETWLLAVLVLLVLLYGFGRTHGLFKKQIGPGKPLPFGTVLVNYMGLWKFV 62  
Qy 55 LSARTGEHAKYREKYGTSLRFAIAGAPVLNSTDPKVFNVHM-KEAY-----DYPKP 106  
Db 63 -----ECKHYGKIKWG---LFDG--QNPLFAITDTTEIKNVLKKECFSTNRRDFGPV 111  
Qy 107 GMAARVLRATGDGVVTAEGEAKHRRIMTSPSAQAQSMVDPIFEKGMELVDKMD 166  
Db 112 GI-----MGKAVSVAKDEWKEKRYALLSPFTTSGLKEMFFIIEQYGDILVKLKQE 163  
Qy 167 AAEKDMVAVGESAGEKATRLT-EGVDVVDKMGVGRATLDVMAAGFDYKSDSLQN----- 219

```

164 A-----ETGKPTMKKVGAGYSDVITSTSTSGVNVDSLNNPKDPFV 204
220 -KTNEL-----YVAFVGLTDGAPTLDLSDFKAINW--DFVPYFRTWKRREIPL 264
205 EKTKLLRFPDFDLFLSVLFPFLT---PIYEMLNICNFPKXDSIEFFK----- 250
265 TQGLAVSRVCIELMEQKQAVLGSDQAVDKKQVQGRDILSLVRANIAANLPESQ-K 323
251 -----KQVYRMKTRLDVSKHRV-----DFLOLMNNAHNSDKKESHTA 290
324 LSDBEVLAQISNLLFAGVETSSVLTWMFHLSEDKAVQDKLREEICQI--DTDMPTLDE 381
291 LSDMEITAQSIIFIFAGVEPTSSLSFVLHSLATHPTQKQLQEEIDRALPNKAPTDT 350
382 LNALPYLEAFVKESLRLDPPSPYANRCLKDEPIPLAEPVIGRDSGVINEVRIKGTWV 441
351 VNEMEYLDVNLNETRLYPIGNRLERVCKDVE-----INGVFMPKGSVV 395
442 MLPLFNINRSKFIVGEDAEERPERPRLVEDVTDSLNS-IEAPYGHQASFISGPRACFGMR 500
396 MIPSVLHSDQHWPE-PEEERPERFSPKNGSGIDPVVLPFG-----NGPNCIGMR 448
501 AVAENKAFVTLRRVQPEP 520
449 ALMMNKLALTKVLQNFSPQ 468

RESULT 15
S50892
Cytochrome P450 3A16 - mouse
N:Contains: oxidoreductase [EC 1.-.-.-]
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 09-May-1997 #text_change 28-Jul-2000
C:Accession: S50892
R:Itouh, S.; Sato, M.; Abe, Y.; Hashimoto, H.; Yanagimoto, T.; Kamataki, T.
Eur. J. Biochem. 226, 877-882, 1994
A:Title: A novel form of mouse cytochrome P450 3A (Cyp3a-16). Its cDNA cloning and expression
A:Reference number: S50892; MUID:95112853; PMID:7813478
A:Accession: S50892
A:Molecule type: mRNA
A:Residues: 1-504 <ITO>
A:Cross-references: EMBL:D26137; NID:G493670; PIDN:BA05133.1; PID:G666968
A:Experimental source: fetal liver
A>Note: In the authors' translation residues 9-20 do not match the nucleotide sequence
A>Note: the authors translated the codon TTC for residue 464 as Glu
C:Genetics:
C:Gene: CYP3a-16
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;
F:303-465/Domain: cytochrome P450 homology <P45>
F:469/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 12.9%; Score 368; DB 2; Length 504;
Best Local Similarity 25.8%; Pred. NO. 2e-16;
Matches 132; Conservative 100; Mismatches 171; Indels 108; Gaps 21;

QY 68 EXYGSTL-RFAGIAGAVLNSTDPKVFNHVM-KEAY-----DYPKPGVAARVLRATG 118
DB 66 EXYKGTWGLFDG--QIPFLVITDPTETIKNLVKECFSVFTNRQDFPFGIMSKSISL--- 120
QY 119 DGWVTAEGEAKHRRIRIIPSLAQAVKSMVPIFLEKGMELVDKQMEDAAEKQNAVGESA 178
DB 121 -----AKDEWKRYRALLSPFTSGNLKEMFPVIEQYG-DILVKYLQEAEGKQFVA--- 171
QY 179 GEKKATRLTEGVYKDWVGRATLDYMALAGFDYKSDSLQNKTNELVYAFVGLTDGFAPT 238
DB 172 -----VKDVLGAYSDVITSTFGVNIIDSLNPNP-----PF 203
QY 239 LDSFKAIM-WDF-----VPYFRWKRRHEIPLTQGLAVSRVRVGIEMEQKQAVLGSA 290
DB 204 VENAKVLRVDFDPLSLVALFPFLTIYEM---LNICMFPKDSIEFFK----FYDRM 256

```

```

291 SDQAVDKQVQGRDILSLVRANIAANLPESQ-LSDBEVLAQISNLLFAGVETSSVLT 349
257 TENELSDKQKHRVDPIYLMWEAYNKSQKSHKALSEIEITAQSIIFIFAGVETSSILS 316
350 WMFHLSEDKAVQDKLREEICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANR 407
317 FTVYSLATHPTDIQKQLQEEIDRALPNKAPTDTVMAMEYLDVNLNETRLYPITNRLQR 376
408 ECLKDEDPIPLAEPVIGRDSGVINEVRIKGTWVMLPLFNINRSKFIVGEDAEERPERPW 467
377 VCKDVE-----INGYIPKGSVTIIPSVLHSDQHWPE-PEEQPERF 420
468 LEDVTDSLNS-IEAPYGHQASFISGPRACFGWRFAVAKAFVTLRRVQPEPIISHPE 526
421 SKENKGSIDPVVLPFG-----NGPNCIGMRFALMMNKLALIKVLQNFSPQC---K 470
527 YEHTILISRPVIGREKEGYQNRLO-VKPV 556
471 ETQIPLKLRS-----ELLQPVAPI 490

```

Search completed: April 2, 2004, 14:01:26  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:59:14 ; Search time 12 Seconds  
(without alignments)  
871.939 Million cell updates/sec

Title: US-10-066-007-1  
Perfect score: 2852  
Sequence: 1 MFILVLLTGLALAAFSWAS.....RIVGREKEGYQNRLOVKPVE 557

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 155783 seqs, 18785033 residues

Total number of hits satisfying chosen parameters: 155783

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA New:  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB	ID	Description
1	370	13.0	504	6	US-10-784-004-427	Sequence 427, App
2	370	13.0	504	6	US-10-784-004-948	Sequence 948, App
3	348.5	12.2	546	6	US-10-275-027A-405	Sequence 405, App
4	347.5	12.2	524	6	US-10-275-027A-877	Sequence 877, App
5	347.5	12.2	524	6	US-10-275-027A-878	Sequence 878, App
6	347.5	12.2	524	6	US-10-275-027A-879	Sequence 879, App
7	345.5	12.1	522	6	US-10-767-701-46006	Sequence 46006, A
8	336	11.8	539	6	US-10-767-701-45390	Sequence 45390, A
9	328	11.5	466	1	PCT-US04-04280-33	Sequence 33, Appl
10	323.5	11.3	568	6	US-10-275-027A-406	Sequence 406, App
11	310.5	10.9	470	6	US-10-776-871-12	Sequence 12, Appl
12	310.5	10.9	496	6	US-10-776-871-10	Sequence 10, Appl
13	310.5	10.9	496	6	US-10-782-695-19	Sequence 19, Appl
14	305.5	10.7	253	6	US-10-767-701-40524	Sequence 40524, A
15	286	10.0	276	6	US-10-767-701-41801	Sequence 41801, A
16	282.5	9.9	374	6	US-10-767-701-45728	Sequence 45728, A
17	281.5	9.9	505	6	US-10-776-871-2	Sequence 2, Appl
18	277.5	9.7	429	6	US-10-275-027A-407	Sequence 407, App
19	272	9.5	368	6	US-10-767-701-45321	Sequence 45321, A
20	257.5	9.0	509	6	US-10-100-683-8244	Sequence 8244, App
21	257.5	9.0	509	6	US-10-152-372-108	Sequence 108, App
22	243.5	8.5	216	6	US-10-767-701-41930	Sequence 41930, A
23	240.5	8.4	209	6	US-10-767-701-39467	Sequence 39467, A
24	238	8.3	413	6	US-10-767-701-43099	Sequence 43099, A
25	228.5	8.0	537	6	US-10-782-695-17	Sequence 17, Appl
26	217.5	7.6	195	6	US-10-767-701-44683	Sequence 44683, A

ALIGNMENTS

RESULT 1

US-10-784-004-427

; Sequence 427, Application US/10784004  
; GENERAL INFORMATION:

; APPLICANT: Biogen Idec

; TITLE OF INVENTION: Surrogate Markers of Pain

; FILE REFERENCE: 08201.6029-00000

; CURRENT APPLICATION NUMBER: US/10784,004

; CURRENT FILING DATE: 2004-02-20

; NUMBER OF SEQ ID NOS: 1251

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 427

; LENGTH: 504

; TYPE: PRT

; ORGANISM: rat

; US-10-784-004-427

Query Match 13.0%; Score 370; DB 6; Length 504;

Best Local Similarity 26.2%; Pred. No. 1.1e-20;

Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY	7	LTGALGAASFWSIASIAFFSLY-LAPRRSLYNLQ-----PNTNYTGNF-LDI 54
Db	3	LLSALTLETWLLAVLVLLYGFGRTHGLFKKQIGPKPLPFFGTVJNYTNGLWKFDV 62
QY	55	LSARTGEEHAKYREKYGSTLRFPAGIAGAPVLNSTDPKFVHYM-KEAY-----DYPKP 106
Db	63	-----ECHKKYKING---LFDG--QMPLFAITDTEMIKNVLVKECFVSFTNRDRFGPV 111
QY	107	GMAARVLRIATGCGVTVTAEGEAHRRHRIIMISLSAQAVKSWPIPLEKGMELVDKMMED 166
Db	112	GI-----MGRAVSVAKDEEWKRYALLSPTFTSGRIKEMFPFIIETQYGDILVKYLKOE 163
QY	157	AAEXDMAVGSAGEKKATRLLET-EGVDVQDWGRATLDVVALAGFDYKSDSLQN----- 219
Db	164	A-----ETGKPTMKVFKGAYSDNDVITSTSGVGVVDSLNPNKDPFV 204
QY	220	-KTNEL-----YVAFVGLTGDGFAPTLDSFKAIMW--DFVPYFRTMKRHEIPL 264
Db	205	EKTKLLRFDFDPLFLSVLFFFLT----PIVEMLNICMFPKDSTIEFFK----- 250
QY	265	TQGLAVSRVRCIELMEQKQAVLGSASDAQVDKKQVGGRDILSLVRAIANALPESQ-K 323
Db	251	-----KPYMKKTRLDSVQKRV-----DFLQMMNHNDSKDKEHSTA 290
QY	324	LSDEVLIAQISNLLFAGYETSSVLTWFMHRLSEDKAVQDKLREEICQI--DTDMPTLDE 381
Db	291	LSDEMIFAQSIIFIFAGYETSSLSFVLHSLATHPTDKLQEEIDRALPNKAPPTYDT 350

Qy	382	LNALPYLEAFVKESLRLOPPSYANRECLKDDEDFIAPVIGRDGSVINEVRITGTWY	441
Dd	351	VMEIMEYLDVLNLTUURLYPIGNRLERLVCKDVE-----INGVFMFKGSVV	395
Qy	442	MPLFLFNINRSKTIYGEDAEPRPERWLEDVTDSLNS-TEAPYHQASFISSGRACFGWRP	500
Dd	396	MISYALHRRDPQHWE-PREEFPERFSKENKGSDIPVYLFFG-----NGPRNCIGMRP	448
Qy	501	AVAEMKAKFLVTLRRVQFEP	520
Dd	449	ALMMMKLALTKVQLNFSQP	468

RESULT 2

```
US-10-784-004-948
; Sequence 948, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 948
; LENGTH: 504
; TYPE: PRT
; ORGANISM: rat
US-10-784-004-948
```

```

RESULT 3
US-10-275-027A-405
; Sequence 405, Application US/10275027A
; GENERAL INFORMATION:
; APPLICANT: Hyaeq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP3/US
; CURRENT APPLICATION NUMBER: US/10/275.027A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 405
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-275-027A-405

```



```

RESULT 6
US-10-275-027A-879
; Sequence 879, Application US/10275027A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP3/US
; CURRENT APPLICATION NUMBER: US/10/275,027A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-275-027A-879

Query Match      12.2%; Score 347.5; DB 6; Length 524;
Best Local Similarity 23.6%; Pred. No. 6.1e-19;
Matches 140; Conservative 104; Mismatches 211; Indels 137; Gaps 26;

QY      3  ILVLITGALGLA-AFSWASIAFFSLYLAPRSSLYNLQGNHTNYFTGNFLDLSARTG- 60
Db      21  LULLVVGWLLARILAWT-----YAFYNCCR--LQCFQPPKRNFWGHLGILTFEEGL 74

QY      61  ----BEHAKYREKYGSTLRFAGIAGAVLNSTDPKVENHMK-EAYDYPKPGMAARLRI 115,
Db      75  KDSQTMSATYSQ--GFTVWLGPi--IPFVLCHPDITRSITNASAAIAPKONLFIPLKP 130

QY      116  ATGDGVVTAEGEAHKRHRIMPSLSAQAVKSVPIFEKGMELVDKMEADAEEKMNAVG 175
Db      131  WLGEIILSGGDKWSRHRRLMTAFHNILKSVITIFNKSANIMLDRKWOHLASE-----G 185

QY      176  BSAGEKKATRLTEBGVDKDWGVGRATLDVNMALAGFDYKSDSLQNKTNLYAVFVGLTDGF 235
Db      186  SSC-----LDMFEHISLMTLDSLQKCIFSDS-HQQRPSF-YIA----- 223

QY      236  APTLDSFKAIN-----WDPVVYFR-----TWKRREHIEPLT 265
Db      224  --TILELSALVKEKSHILOHMFYLYLSDHGRFRHACRLVHDFTDAVIRERRRTLP-T 280

QY      266  QGLAVRRVGIELMEQKKQAVLSGASDAQVDKDKVQGRDILSLVRANIAANLPSOKLS 325
Db      281  QGI-----DDFFKRAKSKTLDFIDVL---LLSKDEDGKALS 314

QY      326  DEVLIAQISNLLFAGYETSTVLTMFHLRSEDKAVQDKLREEICQI---DTDMFTL--DE 381
Db      315  DEDIPAEADTFWFGGHTTASGLSWLYNLARHPEYQERCQVEQLLKORDPKKEIWD 374

QY      382  LNALPYLEAFVKESLRIDPSPVANRECKLDEDFIPLAEPVIGRDGVSINEVRITKTMV 441
Db      375  LAQLPFLTMCVKESLRLHPAPFISRCCTQD-----IVLPDGRV----IPKGIIC 420

QY      442  MLPEFNINRSKFYGEDAAEFRRERWLEDVDTSLNSI-EAPYHQGASFIQSPRACFQWR 500
Db      421  LIDIIGVHNPTW-PPEYVDYDFRP-----DPENSKGRSLAP-IPFSAGFRNICQAF 473

QY      501  AVAEKWAKPLFVTLRRVOFEPIIISHPYEHTLIILSRPRIVGREGVQMRQLQ 552
Db      474  AMAEKVVLAALMLHFLFLDPDTEPR-----RKLEIMRAEGLMLRVE 517

```

```

RESULT 7
US-10-767-701-46006
; Sequence 46006, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cac, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53635)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46006
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2930_1 pep
US-10-767-701-46006

```

```

Query Match      12.1%; Score 345.5; DB 6; Length 522;
Best local similarity 22.9%; Pred. No. 8.6e-19;
Matches 134; Conservative 90; Mismatches 211; Indels 149; Gaps 21;

QY      6  LITGALGIAAQSWSAI-AFFSLYLAPRR-SSLYNLOGPNHNYFTGNFLDILSARTGEH 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      18  LLOGLLAM-LIAWGSYQAATFWLPRRLNRLRAQGLSLEY-----LFFAGDLKEN 69

QY      64  AKYREKYGSTRPAGIAGAPVLANSTDPKVPNHVYKAYDY-----PKPGM---AAR 111
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      70  DRLNDE-----ARSRMSPLSHDIVRVNPHLFTNFKVHGHNISTWFGPIRVITTEAE 122

QY      112  VLRATGD-----GVYTAEGEZHKKRRRIMPISLSAQAVKSWMPI 151
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      123  LTRDVLNKSNGHFEKFIENKRLIKLALGIASHDGEKWAHGERILNPAPHLEKLRMLPA 182

QY      152  FLEKGMEVLVDKQMEDAAEKDMAVGESAGEKKATRLTEGVDKVQWVGRTATDVALAGFD 211
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      183  PSTCSTELIDENENKLAASD-----GSEYLDIC-----210

QY      212  YKSDLQKNKNELKVAFVGLTGDGATLDSFKAIMWDVFPVPTNKKRRHEIPLTQGLAV- 270
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      211  -----PPELHPT-----PGGMIPPPGRSGSLGG-----RTLCLRRGPAGRLTAAL 252

QY      271  -----SRRVGIELMEQKKQAVLG--SASDQAVDKDQVQGRDILSLVLRANI 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      253  HLFFPPGPPPPPHNRRT-----RGRKRLRGWIEKRERAIENTSETSNDDLGLLGSNM 307

QY      315  AANLPESQKLSDEVLQAISNLLFAGYETSSVLTWMFHRLSEDAKVQDKUREICQI-- 372
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      308  DSG--KGSLRSTEDVIEECKLFYFAGMEVTSVLLTWTVLVILGHMPEWODRAREEVLVSFG 366

QY      373  DTDMPITDELNALPYLEAFYKESLSLDDPPSPYANRECLKDBDFIPLAEPVIGRDSVINE 432
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      367  RDKQNFNFDGGLRKLKTVTMILYEVLRLYPPAVSLNRRTFKD-----MTIGG 411

QY      433  VRITKGNVMVLPFLFNINRSKFYIGDEAEFEERPERLWDVDSLNSIER--PYGHQASFIS 490
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      412  ISYFAGVILBPLTVVHHSTDVGKDAHEFKFERFAGISKATDKRPAFFPFG-----465

QY      491  GPRACTGWRFAVAEMKAFIVTLRRVQFE--PIISHPEYEHITL 532

Db      466  GFRICIQNFALLAEAKWALSMTLQRFQOLSPSYTHAPYVTLTL 509
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8  
US-10-767-701-45390  
; Sequence 45390, Application US/10767701  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.





QY 453 FTYGDAEERPERWLEDVDTLSNIRAPYGHQASPTSGPRACFGHRAFAVAKAFLEVT 512  
 Db 378 AW-PDPEVDFR-----DPKIKERSPLAFIPFSGAPRNCIGQAFAMKVVGLIT 431  
 QY 513 LRVQFEPIIHPYEHITLIISRPVGRKEGYQMRLO 552  
 Db 432 LLRFRALPDHTEPR-----RKPELVRAEGGLMLRVE 463

RESULT 10  
 US-10-275-027A-406  
 ; Sequence 406, Application US/10275027A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; APPLICANT: Tang et al.  
 ; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
 ; FILE REFERENCE: 785CIP3/US  
 ; CURRENT APPLICATION NUMBER: US/10/275,027A  
 ; PRIOR FILING DATE: 2002-10-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/02623  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 09/491,404  
 ; PRIOR FILING DATE: 2000-01-25  
 ; PRIOR APPLICATION NUMBER: 09/517,745  
 ; PRIOR FILING DATE: 2000-07-17  
 ; PRIOR APPLICATION NUMBER: 09/631,451  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: 09/633,870  
 ; PRIOR FILING DATE: 2000-09-15  
 ; NUMBER OF SEQ ID NOS: 944  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 406  
 ; LENGTH: 568  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-275-027A-406

Query Match 11.3%; Score 323.5; DB 6; Length 568;  
 Best Local Similarity 22.2%; Pred No. 4.6e-17;  
 Matches 137; Conservative 105; Mismatches 210; Indels 165; Gaps 26;

QY 3 ILVLLTGAALGA-AFWSASIAFSLYLAPRRSSLYNLQGNHTNYFTGNFIDLSARTG- 60  
 Db 43 LLLLVGSMILARILAWT---YAFVNNCR--LQCFPPKPRNFWGHLITPTBEGL 96  
 QY 61 -----EEHAKYREKYGTSLRFAGIAGAPVLNSDTPKVFNVHK-EAYDYKPGMAARVLI 115  
 Db 97 KNTQMSATVQ--GFTIWLGEI--IPFVLCHPDTIRSIITNAGAAIAPKDNLFIRELKP 152  
 QY 116 ATGDGVVTAEGEAHKEHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAG 175  
 Db 153 WLCEGILLSGDKWSRHRMLTPAFHFNILKSYITFNKSNANIMLDKQWHLASE-----G 207  
 QY 176 ESAGEKATRETEGVVDKVDWVGRATLDVMAAGFDYKSDLSQNTKVELYVAFVGLTDGF 235  
 Db 208 SSC-----LDMFEHLSMTLSLQKCIFFSDS-HCQKPSB-YIA----- 245  
 QY 236 APTLDSFKAIM-----WDFEYFER-----TMKRREHIEPLT 265  
 Db 246 --FILELSALVEKRSQHILQHMDFLYLSHDGRRFRACRLVHDFDAVIERRRRTLP-T 302  
 QY 266 QGLAVRRVGIEMEQKQAVGSGASDAQVDKQVQGRDILSLVRANIANLPSQKLS 325  
 Db 303 QGT-----DDFFKQAKAKTKTLDFIDVL-----LLSKDSDGKALS 336  
 QY 326 DEEVLQISNLLF-----AGVETSTVLTMFHRLSDEKAVQD 363  
 Db 337 DEIDRAEDTFMFGGQYLGAVHPVLPKSLPGCSSGHDTTASGLSWLYNLARHPPEQE 396  
 QY 364 KLBREICQI--DTDMPTL--DELNALPYLEAFVKLSRLDDPPSPYANRECUKDEDFPLA 419  
 Db 397 RCQEQVELLKRDPKPEIENWDDLAQLPFLTMCVKLSRLHPPAPFISRCCTQD----- 449

QY 420 EPVICRDSGVINEVRITKGTMTVMPLFNINRSKFIYGE-----DAEEFRPERWLEDVDTSL 475  
 Db 450 --IVLPDGRV-----IPKGITCLIDIIGHVHNPTVMPDGVYDPRFDPDEN----- 493  
 QY 476 NSIEAPYGHQASPTSGPRACFGHRAFAVAKAFLEVTLSRVQFEPIIHPYEHITLIIS 535  
 Db 494 SKGRSPLAF-IPFSGAPRNCIGQAFAMKVVGLMLLHFRFLPDHTEPR-----R 544

QY 536 RPRIVGREKEGYQMRLO 552  
 Db 545 KLEILMEAREGGGLMLRVE 561

RESULT 11  
 US-10-776-871-12  
 ; Sequence 12, Application US/10776871  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glucksmann, Maria Alexandra  
 ; APPLICANT: Meyers, Rachel  
 ; APPLICANT: Rudolph-Owen, Laura A.  
 ; TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF  
 ; FILE REFERENCE: 10448-193001  
 ; CURRENT APPLICATION NUMBER: US/10/776,871  
 ; CURRENT FILING DATE: 2004-02-11  
 ; PRIOR APPLICATION NUMBER: US/10/175,896  
 ; PRIOR FILING DATE: 2002-06-20  
 ; PRIOR APPLICATION NUMBER: 10/067,668  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: 60/286,140  
 ; PRIOR FILING DATE: 2001-02-02  
 ; PRIOR APPLICATION NUMBER: 09/823,901  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/10720  
 ; PRIOR FILING DATE: 2001-04-02  
 ; PRIOR APPLICATION NUMBER: 60/193,920  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/862,658  
 ; PRIOR FILING DATE: 2001-05-21  
 ; PRIOR APPLICATION NUMBER: PCT/US01/16380  
 ; PRIOR FILING DATE: 2001-05-21  
 ; PRIOR APPLICATION NUMBER: 60/205,675  
 ; PRIOR FILING DATE: 2000-05-19  
 ; PRIOR APPLICATION NUMBER: 09/882,837  
 ; PRIOR FILING DATE: 2001-06-15  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 470  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: consensus sequence  
 US-10-776-871-12

Query Match 10.9%; Score 310.5; DB 6; Length 470;  
 Best Local Similarity 24.7%; Pred No. 3.5e-16;  
 Matches 141; Conservative 103; Mismatches 172; Indels 155; Gaps 28;

QY 24 FSLYLAPRRSSLYNLQGNHTNYFTGNFIDLSARTGEEHAKYREKYGTSLRFAGIAGAP 83  
 Db 15 FTYLGLPR--PWVLTGPEAVK-----EVLIDK-GEFEAKGKRGDFNPTF-----P 56  
 QY 84 VLNSDTPKVFNVHMKAEVDYKPGVAAARVLRATGDDGVTAEGEAHKEHRRIMIPSL--- 140  
 Db 57 WLUSK-----GYREQGL-----LFDNGPKWKRLARFSLTLRPH 90  
 QY 141 -----SAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAGSEKAKATRETEGVQVK 194  
 Db 91 FCGVAYSXESQKLEFPRIQEEARDLVERLRK-----EQAGSP-----IDIT 131  
 QY 195 DWVGR-ATLDVMAAGFDYKSDLSQNK-----TNELYVAFVGLTDGFAP---TL 239

Db 132 ELLARLAPLNVICSLFVGVRFDYLRPEDEPEFLKLDKLLNEMF-----DRVSPWHQLL 184  
 Qy 240 DSKAIVMDVFP---VFTWKRHRHEIPTQGLAVSRVGIEMEOKQAVLGSDQAVD 296  
 Db 185 DIFPFL--RYLPGSLFRKAFKAKDL-----KDYLDKLIBERETL-----EPAGD 229  
 Qy 297 KDVQGRDILSLVRANIANLPESOKLSDEEVLQISNLLFAGYETSTVLTWTFHRLS 356  
 Db 230 PRLDIGFDSLILLEAKREGNPKSE--LSDEELAATVLDLLFAGTETTSSTLSWALYLLA 288  
 Qy 357 EDKAVQDKLREETICQI--DTDMPT--LDELNALPYLEAFVKESLRLDPPSPYANRECLD 412  
 Db 289 KHPEVQAKLREEDIVIGDRSRSTYVDARAQMPYLDVAVIKETLRLY- 336  
 Qy 413 EDFTPLAEP--VIGRDSGVINEVRITKGTWMLPLFNINRSKIYGEDAEFPERMLEDV 471  
 Db 337 --VVPFL--RYLPGSLFRKAFKAKDL-----KDYLDKLIBERETL-----EPAGD 255  
 Qy 472 TDSLNSIE--APYGHQASFSISGRACFGRFVAEMKAFVTLRRVQPEPIISHPEYE-- 528  
 Db 394 GKPKKSYAFLPFG-----AGPRNCLGERLARMELFLATLLQR-----PPELELA 439  
 Qy 529 ----HITLISRPRIVGREKEGYQMLQVQP 555  
 Db 440 VPPGDIPSLTPKPEL--GLPSKPLYKVQLRP 469

RESULT 12  
 US-10-776-871-10  
 ; Sequence 10, Application US/10776871  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glucksmann, Maria A.  
 ; APPLICANT: Meyers, Rachel  
 ; APPLICANT: Rudolph-Owen, Laura A.  
 ; TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF  
 ; FILE REFERENCE: 10448-193001  
 ; CURRENT APPLICATION NUMBER: US/10/776,871  
 ; CURRENT FILING DATE: 2004-02-11  
 ; PRIOR APPLICATION NUMBER: US/10/175,696  
 ; PRIOR FILING DATE: 2002-06-20  
 ; PRIOR APPLICATION NUMBER: 10/067,668  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: 60/266,140  
 ; PRIOR FILING DATE: 2001-02-02  
 ; PRIOR APPLICATION NUMBER: 09/823,901  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/10720  
 ; PRIOR FILING DATE: 2001-04-02  
 ; PRIOR APPLICATION NUMBER: 60/193,920  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/862,658  
 ; PRIOR FILING DATE: 2001-05-21  
 ; PRIOR APPLICATION NUMBER: PCT/US01/16380  
 ; PRIOR FILING DATE: 2001-05-21  
 ; PRIOR APPLICATION NUMBER: 60/205,675  
 ; PRIOR FILING DATE: 2000-05-19  
 ; PRIOR APPLICATION NUMBER: 09/882,837  
 ; PRIOR FILING DATE: 2001-06-15  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 496  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: consensus sequence  
 US-10-776-871-10

Query Match 10.9%; Score 310.5; DB 6; Length 496;  
 Best Local Similarity 24.7%; Pred. No. 3.8e-16;  
 Matches 141; Conservative 103; Mismatches 172; Indels 155; Gaps 28;

Qy 24 FSLYLAPRSRLYNLQGNHTNYFTGNFLDILSARTGEEHAKYREKYGSTLRFAGIAGAP 83  
 Db 41 FLYLIGPR--PVVLTGSEAVK-----EVLIDK--GEFAKGRGDFNPIF-----P 82  
 Qy 84 VLNSTDPKVFNHVMKEAVDYKPGMAARVLIATGDGVVTAEGEAHKKHRRIMTSPSL--- 140  
 Db 83 WLSK-----GYREQGL-----LFSNGPKWKRLRRFSLTLTFRFH 116  
 Qy 141 ----SAGAVKSMVPIFLEKGMELVDKMDAAEKOMAVGESAGEKATRLTEGVDDVK 194  
 Db 117 FGMGAYSKSKQLEPRIOEAEADLVERLRK-----EQAGSP-----IDIT 157  
 Qy 195 DWVGR--ATLDVNMALAGFDYKSDSLQNK-----TNELYVAFVGLTQGFAP---TL 239  
 Db 158 ELLARLAPLNVICSLFVGVRFDYLRPEDEPEFLKLDKLLNEMF-----DRVSPWHQLL 210  
 Qy 240 DSKAIVMDVFP---VFTWKRHRHEIPTQGLAVSRVGIEMEOKQAVLGSDQAVD 296  
 Db 211 DIFPFL--RYLPGSLFRKAFKAKDL-----KDYLDKLIBERETL-----EPAGD 255  
 Qy 297 KDVQGRDILSLVRANIANLPESOKLSDEEVLQISNLLFAGYETSTVLTWTFHRLS 356  
 Db 256 PRLDIGFDSLILLEAKREGNPKSE--LSDEELAATVLDLLFAGTETTSSTLSWALYLLA 314  
 Qy 357 EDKAVQDKLREETICQI--DTDMPT--LDELNALPYLEAFVKESLRLDPPSPYANRECLD 412  
 Db 315 KHPEVQAKLREEDIVIGDRSRSTYVDARAQMPYLDVAVIKETLRLY- 362  
 Qy 413 EDFTPLAEP--VIGRDSGVINEVRITKGTWMLPLFNINRSKIYGEDAEFPERMLEDV 471  
 Db 363 --VVPFL--RYLPGSLFRKAFKAKDL-----KDYLDKLIBERETL-----EPAGD 255  
 Qy 472 TDSLNSIE--APYGHQASFSISGRACFGRFVAEMKAFVTLRRVQPEPIISHPEYE-- 528  
 Db 420 GKPKKSYAFLPFG-----AGPRNCLGERLARMELFLATLLQR-----PPELELA 465  
 Qy 529 ----HITLISRPRIVGREKEGYQMLQVQP 555  
 Db 466 VPPGDIPSLTPKPEL--GLPSKPLYKVQLRP 495

RESULT 13  
 US-10-782-695-19  
 ; Sequence 19, Application US/10782695  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glucksmann, Maria A.  
 ; APPLICANT: Curtis, Rory A.J.  
 ; APPLICANT: Tsai, Fong-Yang  
 ; APPLICANT: Hodge, Martin R.  
 ; APPLICANT: Meyers, Rachel E.  
 ; APPLICANT: MacBeth, Kyle J.  
 ; APPLICANT: Bandaru, Rajasekhar  
 ; TITLE OF INVENTION: NOVEL 14275, 54420, 8797, 27439, 68730.  
 ; FILE REFERENCE: 69112 AND 52908 MOLECULES AND USES THEREFOR  
 ; CURRENT APPLICATION NUMBER: US/10/782,695  
 ; CURRENT FILING DATE: 2004-02-19  
 ; PRIOR APPLICATION NUMBER: US 09/945,254  
 ; PRIOR FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: US 60/229,829  
 ; PRIOR FILING DATE: 2000-08-31  
 ; PRIOR APPLICATION NUMBER: US 09/945,301  
 ; PRIOR FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: US 60/229,301  
 ; PRIOR FILING DATE: 2000-09-01  
 ; PRIOR APPLICATION NUMBER: US 10/007,399  
 ; PRIOR FILING DATE: 2001-11-05  
 ; PRIOR APPLICATION NUMBER: US 09/390,039  
 ; PRIOR FILING DATE: 1999-09-03  
 ; PRIOR APPLICATION NUMBER: US 09/146,416  
 ; PRIOR FILING DATE: 1998-09-03  
 ; PRIOR APPLICATION NUMBER: US 10/024,036  
 ; PRIOR FILING DATE: 2001-12-17

```
; PRIOR APPLICATION NUMBER: US 60/258,222
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/103,458
; PRIOR FILING DATE: 2002-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-10-782-695-19

Query Match      10.9%; Score 310.5; DB 6; Length 496;
Best Local Similarity 24.7%; Pred. No. 3.8e-16;
Matches 141; Conservative 103; Mismatches 172; Indels 155; Gaps 28;

QY 24 PSLYAPRSSLYNIQGNHNTYFTGNFLDILSARTGEEHAKYREKYGSTLRFAGIAGAP 83
Db 41 FTLYLGR--PVVLTGPEAVK-----EVLIDK-GEFAKGRGDFNPTF-----P 82

QY 84 VLNSTDPEKVFNVHMKAEVDYDPKGMARVLRIATGCVVTAEGEAHKHRRIMIPSL-- 140
Db 83 WLSK-----GYRQGL-----LPSDNGPKWKLRRFSLLTLRPH 116

QY 141 -----SAQAKSMVPIFEKGMELVDMQMEDAAEKMAVGESAGEKKATRLTEGVGVK 194
Db 117 FGMGAYSKSKLEBPRIQEARDLVERLK-----EQAGSP-----IDIT 157

QY 195 DWVGR-ATLDVNALAGFDYKSDSLQNK-----TNELYVAEVLGTGDPAP---TL 239
Db 158 ELARLAPLVNVCISLLGVRFDYLPEDPEFLKIDKLNNMF-----DRVSPWQLL 210

QY 240 DSFKAIMWDFVP---YFRTKMRHRPIPTQGLAVSRVRVGIEMEQKQKQAVLGASDAOVD 296
Db 211 DIPPEFL-RVLPGLSFRKAFKAKDL-----KDYLDKLIBERRETL-----EPAGD 255

QY 297 KDVQGEDILSLVVRANIANANLPEQKLSDEVLQAISNLLPAGVETSSVLTWFMFRLS 356
Db 256 PRDLTGFDLSLLAEKREGGPKSE-JSDEELATVLDLFPAGTETTSLSWALYLLA 314

QY 357 EDKAVQDKLREETCQT--DITDMP--LDELNALPYLEAFVKESLSRLDPPSPYANRECLD 412
Db 315 KHEVQAKLREEDIVEIGDRSPTYVDVDAQMPYLDVAVIKETLRLYP----- 362

QY 413 EDPIALAP-VIGRDSGVINEVRITKGTWMLPLFNINRSKEFIYGEDAEERPRPRLWEDV 471
Db 363 --VVPILLPRVATKDEIPDGLIPKGTILVINLYSLHRDPKVF-PNPEEPDPERFLDEN 419

QY 472 TDSLSNIE-APYGHQASFTSGPRACFGMRFAVAEMKAFVLTLRVQVPEPIISHPEYE-- 528
Db 420 GKFKSYAFILPFG-----AGRNCLGERLARMELFLFLATLQIR-----PFELELA 465

QY 529 -----HTLIISRPRIVGRSKEGQKMLQKVP 555
Db 466 VPPGDIPSLTPKPEL-GLPSKPKPLYKVLQRP 495

RESULT 14
US-10-767-701-40524
; Sequence 40524, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
```

```
; SEQ ID NO 40524
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C67135_1.pep
US-10-767-701-40524

Query Match      10.7%; Score 305.5; DB 6; Length 253;
Best Local Similarity 32.0%; Pred. No. 3.7e-16;
Matches 85; Conservative 50; Mismatches 102; Indels 29; Gaps 8;

QY 279 MEQKQAVLGASDAQVDKDV--QGRDILSLVVRANIANANLPEQKLSDEVLQAISNL 336
Db 1 LDKQVRGMLMDIKARVANKDTAGYNDLLGLMLEA-CAPEGETPVLSNDEIIIECKTF 59

QY 337 LFAQVETSSVLTWFMFRLSEDKAVQDKLREET-CQIDTDMPTL-DELNALPYLEAFVKE 394
Db 60 FFAGHDITSHLLTWAATLLSTHPEWQDRLEEVRRREGDVEPTRGDALNKLALVNNFLE 119

QY 395 SLRLDPPSPYANRECLDDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKEI 454
Db 120 TRLYGPVSLIQRKAGSDLD-----LGGIRVPSEGAITTIPIATIHDKREV 164

QY 455 YGEDAEERPRPRLWEDVTDLSNIEAPYGHQASFTSGPRACFGMRFAVAEMKAFVLTLR 514
Db 165 WGDGAGEFKPERFENGVTAAKHNA-----LLSFSSGPRSCIGNFAMTEAKAVVAMILQ 220

QY 515 RVQPEPIISHPEYH--ITLIISRPR 538
Db 221 RFALE---LSEKYVHAFMDVITLPR 243

RESULT 15
US-10-767-701-41801
; Sequence 41801, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 41801
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C68995_1.pep
US-10-767-701-41801

Query Match      10.0%; Score 286; DB 6; Length 276;
Best Local Similarity 31.0%; Pred. No. 1.3e-14;
Matches 88; Conservative 46; Mismatches 114; Indels 36; Gaps 8;

QY 277 ELMEQKQAVLGASD--QAVDKDVQGRDILSLVVRANI---AANLPEQKLSDEVL 331
Db 19 QIKNEIESTILLGLISKRMQAMKESAKDDILGLLESNNMRETGENGQSSQGMETDIME 78

QY 332 QISNLLFAGYETSSVLTWFMFRLSEDKAVQDKLRETCQI-DTDMPTLDELNALPYLEA 390
Db 79 ECKLYPFAGMETTSVLLTWTWILLSMPEWQDRAREVLAIFGKQGYDGLSLKTVTM 138

QY 391 FVKESLRDPPSPYANRECLDDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINR 450
Db 139 ILVEVLRYPPAIFASRKYKE-----MVIGDVTYPAGVIVELEFVPMFIH 183

QY 451 SKFIYGEDAEERPRPRLWEDVTDLSNIEA--PYGHQASFTSGPRACFGMRFAVAEMKAF 508
Db 184 DADWGSOVHEFRPRFAEGIAKAKORLAFFPQW-----GPRICIGNFALLEAKMA 237
```



GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 14:00:49 ; Search time 40 Seconds  
(without alignments)  
3553.561 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFTLVLLTGALGLAFAFWSAS.....RIVGREKGYQMRLQKPFVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1070241 seqs, 262374223 residues

Total number of hits satisfying chosen parameters: 1070241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2852	100.0	557	US-10-066-007-1	Sequence 1, Appli
2	2852	100.0	557	US-10-066-007-3	Sequence 3, Appli
3	398.5	14.0	503	US-10-313-963A-56	Sequence 56, Appl
4	396.5	13.9	503	US-10-146-575-2	Sequence 2, Appli
5	385.5	13.5	503	US-09-957-997-3	Sequence 3, Appli
6	384	13.5	537	US-10-425-114-38180	Sequence 38180, A
7	377.5	13.2	527	US-10-425-114-67055	Sequence 67055, A
8	376.5	13.2	547	US-10-425-114-65616	Sequence 65616, A
9	368.5	12.9	547	US-10-425-114-61218	Sequence 61218, A
10	368.5	12.9	560	US-10-425-114-59349	Sequence 59349, A
11	368.5	12.9	562	US-10-425-114-59350	Sequence 59350, A
12	367	12.9	520	US-09-992-901-2	Sequence 2, Appli
13	365.5	12.8	662	US-10-425-114-38827	Sequence 38827, A
14	365	12.8	520	US-10-114-270-60	Sequence 60, Appl
15	353	12.4	524	US-10-424-599-199559	Sequence 199559,

16	352	12.3	533	12	US-10-424-599-269344	Sequence 269344,
17	350	12.3	520	12	US-10-072-012-548	Sequence 548, App
18	350	12.3	520	14	US-10-032-189-115	Sequence 115, App
19	349.5	12.3	520	12	US-10-424-599-228819	Sequence 228819,
20	349	12.2	508	14	US-10-274-694-16	Sequence 16, Appl
21	348.5	12.2	546	15	US-10-291-265-405	Sequence 405, App
22	347.5	12.2	524	15	US-10-291-265-877	Sequence 877, App
23	347.5	12.2	524	15	US-10-291-265-878	Sequence 878, App
24	347.5	12.2	524	15	US-10-291-265-879	Sequence 879, App
25	346.5	12.1	524	9	US-09-989-722-264	Sequence 264, App
26	346.5	12.1	524	9	US-09-989-723-264	Sequence 264, App
27	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
28	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
29	346.5	12.1	524	9	US-09-989-731-264	Sequence 264, App
30	346.5	12.1	524	9	US-09-989-732-264	Sequence 264, App
31	346.5	12.1	524	9	US-09-991-073-264	Sequence 264, App
32	346.5	12.1	524	9	US-09-990-442-264	Sequence 264, App
33	346.5	12.1	524	9	US-09-981-163-264	Sequence 264, App
34	346.5	12.1	524	9	US-09-983-604-264	Sequence 264, App
35	346.5	12.1	524	9	US-09-990-456-264	Sequence 264, App
36	346.5	12.1	524	9	US-09-989-721-264	Sequence 264, App
37	346.5	12.1	524	9	US-09-982-598-264	Sequence 264, App
38	346.5	12.1	524	9	US-09-989-293A-264	Sequence 264, App
39	346.5	12.1	524	9	US-09-989-735-264	Sequence 264, App
40	346.5	12.1	524	9	US-09-990-444-264	Sequence 264, App
41	346.5	12.1	524	9	US-09-991-181-264	Sequence 264, App
42	346.5	12.1	524	9	US-09-989-730-264	Sequence 264, App
43	346.5	12.1	524	9	US-09-990-436-264	Sequence 264, App
44	346.5	12.1	524	9	US-09-993-687-264	Sequence 264, App
45	346.5	12.1	524	10	US-09-989-734-264	Sequence 264, App

## ALIGNMENTS

### RESULT 1

US-10-066-007-1  
Sequence 1, Application US/10066007  
Publication No. US2003007691A1  
GENERAL INFORMATION:  
APPLICANT: HOSHINO, Tatsuo  
APPLICANT: OJIMA, Kazuyuki  
APPLICANT: SETOGUCHI, Yutaka  
TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
CURRENT APPLICATION NUMBER: US/10/066,007  
CURRENT FILING DATE: 2001-02-01  
PRIOR APPLICATION NUMBER: US/09/518,386  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: EP 99104668.1  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: EP 00101666.6  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 557  
TYPE: PRT  
ORGANISM: Phaffia rhodozyma  
FEATURE:  
NAME/KEY: TRANSIT  
LOCATION: (1)..(26)  
US-10-066-007-1

Query Match 100.0%; Score 2852; DB 14; Length 557;

Best Local Similarity 100.0%; Pred. No. 2.3e-248; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFTLVLLTGALGLAFAFWSASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDLSARTG 60

Db 1 MFTLVLLTGALGLAFAFWSASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDLSARTG 60

Qy 61 EEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFHVMKEAYDYPKPGMAARVLRIATGDG 120

```
Db 61 EEHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRATG 120
Qy 121 VVTAGEAHRHRRIMISLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKDVAVGESAGE 180
Db 121 VVTAGEAHRHRRIMISLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKDVAVGESAGE 180
Qy 181 KXATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELNLYVAFVGLTGGFAPTL 240
Db 181 KXATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELNLYVAFVGLTGGFAPTL 240
Qy 241 SFKAIMDVPYFRTMKRRHEIPLTQGLAVSRVGIEMEQKQAVLGSASDAQVDDKDV 300
Db 241 SFKAIMDVPYFRTMKRRHEIPLTQGLAVSRVGIEMEQKQAVLGSASDAQVDDKDV 300
Qy 241 SFKAIMDVPYFRTMKRRHEIPLTQGLAVSRVGIEMEQKQAVLGSASDAQVDDKDV 300
Db 301 QGRDILSLLVRANIAANLPSQKLSDEEVLQAINLLFAGYETSSVTLTWFMFRLSEDKA 360
Qy 301 QGRDILSLLVRANIAANLPSQKLSDEEVLQAINLLFAGYETSSVTLTWFMFRLSEDKA 360
Db 301 QGRDILSLLVRANIAANLPSQKLSDEEVLQAINLLFAGYETSSVTLTWFMFRLSEDKA 360
Qy 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLRLDPPSPYANRECLKDEDFTPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLRLDPPSPYANRECLKDEDFTPLAE 420
Qy 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFYIGEDAEEPRERWLEDVTDLSNLSIEA 480
Db 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFYIGEDAEEPRERWLEDVTDLSNLSIEA 480
Qy 481 PYGHOASFISSGRACFGWRFAVAKMFLVTLRRVQFEPISHPYEYHITLIIISPRIV 540
Db 481 PYGHOASFISSGRACFGWRFAVAKMFLVTLRRVQFEPISHPYEYHITLIIISPRIV 540
Qy 541 GREKEGYQMLQVKPVE 557
Db 541 GREKEGYQMLQVKPVE 557
```

## RESULT 2

```
US-10-066-007-3
; Sequence 3, Application US/10066007
; Publication No. US20030077691A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US/09/518,386
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-10-066-007-3
```

```
Query Match 100.0%; Score 2852; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.3e-248;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFTLVLTGALGLAAPSWSIASIAFFSLYLAPRSSLYNLOGNHNTHYFTGNFELDILSARTG 60
Db 1 MFTLVLTGALGLAAPSWSIASIAFFSLYLAPRSSLYNLOGNHNTHYFTGNFELDILSARTG 60
Qy 61 EEHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRATG 120
Db 61 EEHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRATG 120
```

```
Qy 121 VVTAGEAHRHRRIMISLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKDVAVGESAGE 180
Db 121 VVTAGEAHRHRRIMISLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKDVAVGESAGE 180
Qy 181 KXATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELNLYVAFVGLTGGFAPTL 240
Db 181 KXATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELNLYVAFVGLTGGFAPTL 240
Qy 241 SFKAIMDVPYFRTMKRRHEIPLTQGLAVSRVGIEMEQKQAVLGSASDAQVDDKDV 300
Db 241 SFKAIMDVPYFRTMKRRHEIPLTQGLAVSRVGIEMEQKQAVLGSASDAQVDDKDV 300
Qy 301 QGRDILSLLVRANIAANLPSQKLSDEEVLQAINLLFAGYETSSVTLTWFMFRLSEDKA 360
Db 301 QGRDILSLLVRANIAANLPSQKLSDEEVLQAINLLFAGYETSSVTLTWFMFRLSEDKA 360
Qy 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLRLDPPSPYANRECLKDEDFTPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLRLDPPSPYANRECLKDEDFTPLAE 420
Qy 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFYIGEDAEEPRERWLEDVTDLSNLSIEA 480
Db 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFYIGEDAEEPRERWLEDVTDLSNLSIEA 480
Qy 481 PYGHOASFISSGRACFGWRFAVAKMFLVTLRRVQFEPISHPYEYHITLIIISPRIV 540
Db 481 PYGHOASFISSGRACFGWRFAVAKMFLVTLRRVQFEPISHPYEYHITLIIISPRIV 540
Qy 541 GREKEGYQMLQVKPVE 557
Db 541 GREKEGYQMLQVKPVE 557
```

## RESULT 3

```
US-10-313-963A-56
; Sequence 56, Application US/10313963A
; Publication No. US20040002078A1
; GENERAL INFORMATION:
; APPLICANT: Boutell, Jonathan
; APPLICANT: Godber, Benjamin
; APPLICANT: Hart, Darren
; APPLICANT: Blackburn, Jonathan
; TITLE OF INVENTION: Arrays
; FILE REFERENCE: KIL-001
; CURRENT APPLICATION NUMBER: US/10/313,963A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/335,806
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/410,815
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 56
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-313-963A-56
```

```
Query Match 14.0%; Score 398.5; DB 15; Length 503;
Best Local Similarity 28.7%; Pred. No. 6.2e-27;
Matches 153; Conservative 98; Mismatches 195; Indels 97; Gaps 24;

Qy 13 LAAPSW--ASTAFPSLYLAPRRS--SLYNLOG-PNHTNY-FTGNFELDILSARTG-----E 61
Db 7 LAMETWLLLAVALSVLLVLYGTHSHGLFKLGIPOPTPLPFLGN---ILSYHKGFCDME 63
Qy 62 BHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRAT 117
Db 64 CHKYGKVGWGF--YDG--QQPVLAITDPMIKTVLVKCYSVFTNRRPFGPVFKSA- 117
Qy 118 GDGVVTAEGEAHRHRRIMISLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKDVAVGES 177
```

```

Db 118 ----ISIADEEWKRLSLLSPFTSGKLEMPPIIAQYGDVULVRLREA----- 164
QY 178 AGEKKAATRLT-EGVDVKDWGRATLDVVALAGFYKSDSLQNKTNELYVAFVGLTDGFA 236
Db 165 -----ETGKPVTLKDVFGAYSDVITSTSGVNDLSLNNPD----- 201
QY 237 PTLDSFKAIM-WDFV-PYRTMK-RRHEIPLTQGLAV---SRVGIEMEQKQAVLGS 290
Db 202 PFVENTKLLRFDLDFLFLSIIVFFFLIPIILEVINICVFFPREVTFNRKSVKR-----M 256
QY 291 SDQAVDKVQGRDILSLVLRANIANLPESOKLSDEEVLAQISNLLFAGYETSTVLTW 350
Db 257 KESRLDTQGRVDFLQIMDSQSKTESHKALSDLELVAQSIIFIFAGYETTSVLSF 316
QY 351 MFHRLSEDKAVQDKLREBEICQI---DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408
Db 317 IMVELATHPDVQOKLQEBEIDAVLPNKAPPTYDTVLQMEYLDVMMVNETLRLFPFIAMRLERV 376
QY 409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTMMVMLPLFNINRSKFIYGEDAEERPERWL 468
Db 377 CKXKVE-----INGMFIPKGWVWIPSYALHRDP-KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWFAVAEMKAPFLFVTLRRVQPEP 520
Db 421 KKKKNIDPIYITPFG-----SGPRNCIGMRFPALNMKMLALIRVLQNFSPKP 467

```

## RESULT 4

US-10-146-575-2

; Sequence 2, Application US/10146575

; Publication No. US20030059800A1

; GENERAL INFORMATION:

; APPLICANT: Lichter, Jay

; APPLICANT: Guido, Marco

; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4

; CURRENT APPLICATION NUMBER: US/10/146,575

; CURRENT FILING DATE: 2002-05-14

; PRIOR APPLICATION NUMBER: US/09/144,367

; PRIOR FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 503

; TYPE: PRT

; ORGANISM: H. sapiens

US-10-146-575-2

```

Query Match      13.9%; Score 396.5; DB 14; Length 503;
Best Local Similarity 28.7%; Pred. No. 9.4e-27;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

```

```

QY 13 LAAFSW--ASIAFFSYLAPRRG-SLYNLQ--PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLAVSLVLYLYGTHSHGLFKKLGIPGPTPLPFLGN---ILSYHKGFCWFDME 63
QY 62 EHAKYREKYGSTLRFAGIAGAPVNLSTDPKVFNVHM-KEAYD---PKPGMAARVLRAT 117
Db 64 CHKYGKVGWFG---YDG--QQPVLATDPMIKTVLVKCYSVFTNRRFPFGVGFMSA- 117
QY 118 GDGVVTAEGEAHKKHRRIMIPISLAQAVKSMVPIFLEKGMELVDKXMEDAAEKOMAVGES 177
Db 118 ---ISIADEEWKRLSLLSPFTSGKLEMPPIIAQYGDVULVRLREA----- 164
QY 178 AGEKKAATRLT-EGVDVKDWGRATLDVVALAGFYKSDSLQNKTNELYVAFVGLTDGFA 236
Db 165 -----ETGKPVTLKDVFGAYSDVITSTSGVNDLSLNNPD----- 201
QY 237 PTLDSFKAIM-WDFV-PYRTMK-RRHEIPLTQGLAV---SRVGIEMEQKQAVLGS 290
Db 202 PFVENTKLLRFDLDFLFLSIIVFFFLIPIILEVINICVFFPREVTFNRKSVKR-----M 256
QY 291 SDQAVDKVQGRDILSLVLRANIANLPESOKLSDEEVLAQISNLLFAGYETSTVLTW 350

```

```

Db 257 KESRLDTQGRVDFLQIMDSQSKTESHKALSDLELVAQSIIFIFAGYETTSVLSF 316
QY 351 MFHRLSEDKAVQDKLREBEICQI---DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408
Db 317 IMVELATHPDVQOKLQEBEIDAVLPNKAPPTYDTVLQMEYLDVMMVNETLRLFPFIAMRLERV 376
QY 409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTMMVMLPLFNINRSKFIYGEDAEERPERWL 468
Db 377 CKXKVE-----INGMFIPKGWVWIPSYALHRDP-KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWFAVAEMKAPFLFVTLRRVQPEP 520
Db 421 KKKKNIDPIYITPFG-----SGPRNCIGMRFPALNMKMLALIRVLQNFSPKP 467

```

## RESULT 5

US-09-957-997-3

; Sequence 3, Application US/09957997

; Patent No. US20020150915A1

; GENERAL INFORMATION:

; APPLICANT: Berkenstam, Anders

; APPLICANT: Bertilsson, Gran

; APPLICANT: Blomquist, Patrik

; TITLE OF INVENTION: PROMOTER SEQUENCES

; FILE REFERENCE: 13425-046001

; CURRENT APPLICATION NUMBER: US/09/957,997

; EARLIER FILING DATE: 2001-09-21

; EARLIER APPLICATION NUMBER: SE 0003393-6

; EARLIER FILING DATE: 2000-09-22

; EARLIER APPLICATION NUMBER: 60/238,895

; EARLIER FILING DATE: 2000-10-10

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 503

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-957-997-3

```

Query Match      13.5%; Score 385.5; DB 9; Length 503;
Best Local Similarity 29.0%; Pred. No. 9.3e-26;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

```

```

QY 13 LAAFSW--ASIAFFSYL-APRESSLYNLQ--PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAVETWLLAVSLVLYLYGTRTHGLFKKLGIPGPTPLPFLGN---ALSPKGYWTFDME 63
QY 62 EHAKYREKYGSTLRFAGIAGAPVNLSTDPKVFNVHM-KEAYD-----YKPGMAARVL 113
Db 64 CYKRYKVGW-----YDQQQPMALITDPMIKTVLVKCYSVFTNRRFPFGVGFMSKNAI 118
QY 114 RIATGDGVVTAEGEAHKKHRRIMIPISLAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMA 173
Db 119 SI-----ADEDEWKRIIRLSLSPFTSGKLEKMPPIIAQYGDVULVRLREA----- 164
QY 174 VGESAGEKKAATLET-EGVDVKDWGRATLDVVALAGFYKSDSLQNKTNELYVAFVGLT 232
Db 165 -----ETGKPVTLKGVFGAYSDVITSTSGVSDLSLNNPD-----PFVENT 207
QY 233 D---GFAPTLDSFKAIMWDFVPIFTMKRRHEIPLTQGLAVS---RRVGIEMEQKQAV 286
Db 208 KKLRENP-LDPFVLSIKVP-PFL-----TPILEALNITVFPKVISFLTSKVKQIK 257
QY 287 LGSASQAVDKKDVGR--DILSLVLRANIANLPESOKLSDEEVLAQISNLLFAGYETS 344
Db 258 EGEL-----KETQKGRVDFLQIMDSQSKTESHKALSDLELVAQSIIFIFAGYETT 310
QY 345 STVLTWFMHRLSEDKAVQDKLREBEICQIDTM-----FTLDELNALPYLEAFVKESLRD 399
Db 311 SSVLSFIIYELATHPDVQOKVQKE---IDTLPNKAPPTYDTVLQLELDVMMVNETLRLF 367
QY 400 PPSPYANRECLNDEDFIPLAEPVIGRDSGVINEVRITKGTMMVMLPLFNINRSKFIYGEDA 459

```



Db 368 PYAMRLERCKDVE-----INGMFIPKGVVMPISVVLHDP-KYWTEP 411  
QY 460 EFRPRERWLEDVDSLNS-IEAPYCHOASFSIGRACFGWEPFAVAEMKAEFLVTLRQVF 518  
Db 412 EKFLPERFSKKNKNDIPVITPFG-----SGPRNCIGMRFALVNNKIALVRVLQNFSP 465  
QY 519 EP 520  
Db 466 KP 467  
RESULT 6  
US-10-425-114-38180  
; Sequence 38180, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 38180  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3608-017-B5\_FLI\_1.p  
US-10-425-114-38180

Query Match 13.5%; Score 384; DB 12; Length 537;  
Best Local Similarity 27.4%; Pred. No. 1.4e-25;  
Matches 162; Conservative 79; Mismatches 230; Indels 120; Gaps 24;  
QY 4 LVLLTGALGAAPSWASIAFFS-LYIAPRR-SSLNVLQGNHNTY--FTGNFLDILSART 59  
Db 23 LLIVYGVLG-ALLLWKAARLLRLMWEPRRLERALLAQLRGTSYRFLTG---DLREYRR 78  
QY 60 GEEHAKYRE-----KYGSTLRFAGIAGAPVLNSTDPKVFNVMK 98  
Db 79 SKEENARPLRCHDIAGHVEFFIHGAVLEHGKTC-FTWFGVPVRVTVDPDLARDVMA 137  
QY 99 EAY-DYKPGMAARVLIATGDGVVTAEGEAHKKRRIRIMPSLSAQAVKSMVPIFLEKGM 157  
Db 138 NKTGHPEKPPA--LTKLPDGVANHEGKWKVKKRRILNPAFLKELKMLPAFSACCE 195  
QY 158 ELVDKMEADAEDKMAVGSAGEKKATRLTEGVVDKQVKGWGRATLDVMAAGF-----D 211  
Db 196 ELVSRW-----AOSLGDGCEI-----DVEDELQTLTGDVVISRTAFGSSVLEG 239  
QY 212 YKSDSLQNKTNELVAFVGLTDFAPTLDSFKAIMWDFVPYFRMTKRHEIPLTQGLAVS 271  
Db 240 RKIFOLQAEQERLMSII--DKFA-----VPGVMSLPTKNNRMRQIKSEI 283  
QY 272 RRVGIELMEQKQAVLGASDAQVDKQVGRDILSLVRANIANLPES-----QKLSDE 327  
Db 284 DSILRGLIGKRMQAMKQGESD-----KD-----DLGLLLESNARETGDSGQPGGLTME 334  
QY 328 EVLAQISNLFLFAGVETSSVLTWTFHRLSEDKAVQDKLREECQI--DTDMPTLDLNLAL 385  
Db 335 EWMEECKLFVAGNETSVLLTWTWVLLSWHPEWQDBAREEVLGFGKKQPGYDGLSRL 394  
QY 386 PYLEAFVKESRLRDPSPVANRECLDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPL 445  
Db 395 KTVTMILYEVLRLYPFAIFSRTYKE-----MVVGDVTPYAGVTLLELV 439

QY 446 FNINRSKIYGEDAEAEPRPERWLEDVDSLNSIEA--PYGHOASFSIGRACFGWRFVA 503  
Db 440 LFIHHPDIIWGSDAHEPRPERFAEGVARASKORLAFPPFCW-----GPRICIGONFALL 493  
QY 504 EMKAFIVTVRRVQFEPILISHPEYHTLILSRPR--IVGREKEGYQMLQ 552  
Db 494 EAKWALSMLQRFOFE---LAPITYTHV-----PRRVIMLRPMHGGAQIKLR 535  
RESULT 7  
US-10-425-114-67055  
; Sequence 67055, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 67055  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4757-032-B12\_FLI\_1.p  
US-10-425-114-67055  
Query Match 13.2%; Score 377.5; DB 12; Length 527;  
Best Local Similarity 24.9%; Pred. No. 5.3e-25;  
Matches 141; Conservative 106; Mismatches 212; Indels 107; Gaps 25;  
QY 6 LLTGAALGAAPSWASIAFFSL---YLAPRRSSLYNLQGNHNTYFTGNFLDILSARTG-- 60  
Db 24 LLIGAL-----FFLVWKYVVVTRWFRGGGIGPSY-RFLVGLSLPEIKRMAAGS 71  
QY 61 -----BEHAKYREKIGSTLRFAGIAGAPVLNSTDPKVFNVHMKAYD-Y 103  
Db 72 KITLDVGDHDFVPIVQPYRRWVSVDYKTFLY-WFGAVPTICVAEVLGVKQVLAETGLF 130  
QY 104 PKPGMAARVLIATGDGVVTAEGEAHKKRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKM 163  
Db 131 PKDYNDIS-MEVLGKGLVLANGEDWKRHEVHVHPAFKPKDKLTMSVWM---ADLVQRM 185  
QY 164 MEDAAEKDMVGSAGEKKATRLTEGVVDKQVKGWGRATLDVMAAGF--DYKSDSLQNK 221  
Db 186 MQQWRSQ-----IQRASNHEAE-IELSSEFSELTSDVIAHTAFGTSYK-----EG 229  
QY 222 NELVAVFGLTGDGAPTLDSFKAIMWDFVPYFRMTKRHEIPLTQGLAVSRVRVGIEMEQ 281  
Db 230 KEVFA-----QKELQELTFTWLDIPAPACL-RKLKLPSTKS---SERV--EELDK 275  
QY 282 KQAVLGASDAQVDKQVQ--GRDILSLVRANIANLPESQKLSDEEVLQISNLFLPA 339  
Db 276 KVRSLMAIIEGRLAARCTSGYGNLGLMLQAR-ALEQEGHQLMTTEEVDSCKTFIFA 334  
QY 340 GYETSSTVLTWTFHRLSEDKAVQDKLREECQIDTD-MPTLDLNLALPYLEAFVKESLRL 398  
Db 335 GQDTTSHLTLTMTFLLSRYSEWQRLREEVURECGDAVPNPDTVTKLKNVWVLLSRL 394  
QY 399 DPSPVANRECLDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKIYGED 458  
Db 395 YSPVVFIR-----AVGSD--ILLSTRVPRKGTMTLSIPTIALLRDKVWGQD 439  
QY 459 ABEFPERWLEDVDSL-----NSTEAPYCHOASFSIGRACFGWRFVAEMKAEFLVTLR 514  
Db 440 ADEFNDRPFEGHVSNAAKHPNAL-----LSFSGQFRACIGONFANLEARIGIAMILQ 492

Db 500 LLEAKWTLCTILQRFSELSFSPSYTHAPYTVITL-----HPQGAQIRLK 543

RESULT 9

US-10-425-114-61218

; Sequence 61218, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 61218

; LENGTH: 547

; TYPE: PRT

; ORGANISM: Oryza sativa nipponbare

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3431-054-B3\_FLI pep

US-10-425-114-61218

Query Match 12.9%; Score 368.5; DB 12; Length 547;

Best Local Similarity 24.6%; Pred. No. 3.6e-24;

Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

Qy 4 LVLLTGALGLAFAFASW--ASTAFFSLYLAPRSLLYLOGPNHTNY--FTGNF-----51

Db 36 ILAATAAAVAALLMLAVSTLEWAWTPRLRALRAQGRNRYRLFTGDPENVRNR 95

Qy 52 -----LDILS-----ARTGEHAKYREKYSTLRPAAGAPVLNSDTPKVFN 94

Db 96 EARKKPLGCHDIIIPRLPMFSKAVEHGK-----PSFTWFGPTPRVMISDPESIR 147

Qy 95 HVNKEA---YDYPKPMARVRIATGQVVTAEAGAHKRRHRIIMPISLSAQAVKSVPI 151

Db 148 EVNSNFGHYGKPKRLGLKL-----ASGVSYEGEKWAKHRRILNPAFHHEKIKRMPLV 203

Qy 152 FLEKGMELVDKMDAAEKDMAGVSGAGEKKATRLTEGVDYKDWVGRATLDVNMALAGF- 210

Db 204 FSNCTETWTRW-----ENSMIS-EGMSB-----VDWPEFQNLTDGVISKTAFG 247

Qy 211 -DYKSD-----SLQNKTNELYVAFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPLT 265

Db 248 SSVEEGRRIPLQAESAEIRII-----QAFRTI-----FTPGYWFPLTKNNRLR 291

Qy 266 QGLAVSRVGIEMEQKQAVLGSASDAQVKKDQGRDILSLVRANI-AANLPESQKL 324

Db 292 E-----IREVSKLL-----RGITGK-RERAIKNGETSGDILLGLLVESNWSNGKALGM 342

Qy 325 SDEVLQAIQSNLIPAGYETSVTLTMFHRLSKAVQDKLREEIC-QIDTDMPTLDELN 383

Db 343 TTDIEIECKLFYFAGMETTSVLLTWTLVLSMHPWQERAREEVLHFGRTTPDYDSL 402

Qy 384 ALPYLEAFVKESIRLDPSPFANRECKDEDFIPLAEPVIGRDSVINEVRIKGTVMWL 443

Db 403 RLKVTVMILYEVLRPLPPVVFVFLTRTRYKEMEL-----GGIKYPAEVI-----LML 447

Qy 444 PLFINRSKFIYGEDAEERPERPWLDEVDTSLSNIEAPYGHQASFTIS---GPRACFGWR 500

Db 448 PILFIHDDPDKGDAFENFGFADGINSATK-----YQTSFPFGWGPICIGQNF 500

Qy 501 AVAEKMAFLVTLRRVQFPEIISHPEYEH---ITLISRPRIVGREKEGYQMLQ 552

Db 501 ALLEAKMAICTILQRFSEF-----LSPSYIHAPFTVITLHP-----QHGAQIKLK 545

Db 515 RVQFE--PIISHPEYEHITLIISRPR 538

Db 493 RFSFELSPNVHAPKAEAVILM---PR 515

RESULT 8

US-10-425-114-65616

; Sequence 65616, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 65616

; LENGTH: 547

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700100628\_FLI pep

US-10-425-114-65616

Query Match 13.2%; Score 376.5; DB 12; Length 547;

Best Local Similarity 24.3%; Pred. No. 6.9e-25;

Matches 144; Conservative 96; Mismatches 220; Indels 133; Gaps 25;

Qy 7 LTGALGLAFAFASW--ASTAFFSLYLAPRSLLYLOGPNHTNY--FTGNF-----62

Db 37 LAGAVASVLLNLVAVTLEWAWTPRLRALRAQGRNRYRLFTGDLRE--TARVNR 94

Qy 63 -----HAKYREKYSTLRPAAGAPVLNSDTPKVFNHVKEAYD 102

Db 95 ARKKPLGCHDIIIPRYQPMHSHIKEYGK-LSFTWFGPTPRVMIIPDPELVKELSNKFG 153

Qy 103 Y---PKPCMAARVRIATGQVVTAEAGAHKRRHRIIMPISLSAQAVKSVPIFLEKGMEL 159

Db 154 HFGKPRSRIRGLKL-----ANGLNVHDGKWKAKHRRILNPAFHHEKIKGMWPFSTCCIEW 209

Qy 160 VDKWMDAAEKDMAGVSGAGEKKATRLTEGVDYKDWVGRATLDVNMALAGF--DYKSD-- 215

Db 210 ITRW-----DNSMS-SEGSSB-----IDWPEFQNLTDGVISRTAFGSNYQEGRR 253

Qy 216 --SLQNKTNELYVAFVGLTDGFAPTLDSFKAIMWDFVP--YFRTMKRHEIPLTQGLAYS 271

Db 254 IFELQGEALERLI-----QSVQTFIPGYWFLPTKNNRRM-----288

Qy 272 RVVGLIEMEQKQAVLGSASDAQVKKDQGRDILSLVRANI-AANLPESQKLS 325

Db 289 RAIDVEIKRIRE-IIGKR-----EKDKNRETNKDDLGLLLESNTRSGNAGSLGT 341

Qy 326 DEEVLQAIQSNLIPAGYETSVTLTMFHRLSKAVQDKLREEI-QIDTDMPTLDELNA 384

Db 342 TEDVIECKLFYFAGMETTSVLLTWTLVLSMHPWQERAREEVLHFGRTTPDYDSLGR 401

Qy 385 LPYLEAFVKESIRLDPSPFANRECKDEDFIPLAEPVIGRDSVINEVRIKGTVMWL 444

Db 402 LKVTVMILYEVLRPLPPVVFVFLTRTRYKEME-----LGGIKYPAEVI-----446

Qy 445 LENINRSKFIYGEDAEERPERPWLDEVDTSLSNIEAPYGHQASFI---SQPRACFGWRFA 501

Db 447 VIFIHDDPDKGDAFENFGFADGINSATK-----ANGISSATRHOAAPPFGGPGICIGQSF 499

Qy 502 VAEMKAFVTLRRVQFPEIISHPEYEHITLIISRPRIVGREKEGYQMLQ 552

RESULT 10  
 US-10-425-114-59349  
 ; Sequence 59349, Application US/10425114  
 ; Publication No. US2004003488A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 59349  
 ; LENGTH: 560  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa nipponbare  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: JC-OSLELIB3474019G01\_FLI.pep  
 ; US-10-425-114-59349

Query Match 12.9%; Score 368.5; DB 12; Length 560;  
 Best Local Similarity 24.6%; Pred. No. 3.8e-24;  
 Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLLTGALGLAFAFWSIASIAFFS-LYLAPRR-SSLYNLOGPNHTNY--FTGNF-----51  
 Db 49 ILAAAAAVALLLWAVSTLWAWTPRRLERALARAGIRGNRYRLTGDVPENVRNLNR 108  
 QY 52 -----LDILS-----ARTGEEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFN 94  
 Db 109 EARKKPLPLGCHDIIIPVLPMFSKAVEHGK-----PGTFWGPTRVWISDESIR 160  
 QY 95 HVMKEA---YDYPKPGMAARVLIATGDGVVTAEGEAHKHRRIMIPSLSAQAVKSMVPI 151  
 Db 161 EVMSNKGHYGKPKPTRLGKLL-----ASGVVSVEGEKWKAKHRRILNPAFHHEKIKRMLPV 216  
 QY 152 FLEKGMELVDKMDAEDAAEKMAVGESAGEKATKATLETEGVDVVDKDWVGRATLDVWALAGF- 210  
 Db 217 FSNCTETWTRW-----ENSMI-EGMSE-----VDVWPEFQNLTDGVISKTAFG 260  
 QY 211 -DYKSD-----SLQNKTNELYVAFGLTDGFAPTLDSFKAIMWDFVPYFRMKRRHEIPT 265  
 Db 261 SSYEGRRIIFQLOAESAEIRII-----QAFRTI-----FIPGYWFLPTKNNRRLR 304  
 QY 266 QGLAVSRVGIEMLEQKQAVLGSASDAQVDKQVGRDILSLVRANI-AANLPESOKL 324  
 Db 305 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLDGLLVESNMRESNGKAEI 355  
 QY 325 SDEVLAQISNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREEIC-QIDTDMPTLDELN 383  
 Db 356 TTDEIIIECKLFYFAGMETTSVLLTTLVLSMHPWQERAREEVLHFGRTTPDYDSLS 415  
 QY 384 ALPYLEAFVKESLRDPPSPYANRECKLDEDFIPLAEPVIGRDSGVINEVRITKGTWML 443  
 Db 416 RLKIVTMLIYEVRLPVVFLTRTYKWEML-----GGIKYPAEVT-----LML 460  
 QY 444 PLFNINRSKFIYGEDAEERFRERWLEDVDTSLNSIEAPYGHQASFTIS---GPRACFGWR 500  
 Db 461 PILFIHDDPDWKGADGEFNGRPFADGISNATK-----YQTSFPFGWGPICIGQNF 513  
 QY 501 AVAEMKAFVTLRRVQPEPIIISHPEYH---ITLIISRPRIVGREKEGYQWRLO 552  
 Db 514 ALLEAKMAICTILQRFSE---LSPSYTHAPFTVITLHP-----OHGAQIKLK 558

RESULT 11  
 US-10-425-114-59350  
 ; Sequence 59350, Application US/10425114

Publication No. US2004003488A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 59350  
 ; LENGTH: 562  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-OSFLCYP143F07\_FLI.pep  
 ; US-10-425-114-59350

Query Match 12.9%; Score 368.5; DB 12; Length 562;  
 Best Local Similarity 24.6%; Pred. No. 3.8e-24;  
 Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLLTGALGLAFAFWSIASIAFFS-LYLAPRR-SSLYNLOGPNHTNY--FTGNF-----51  
 Db 51 ILAAAAAVALLLWAVSTLWAWTPRRLERALARAGIRGNRYRLTGDVPENVRNLNR 110  
 QY 52 -----LDILS-----ARTGEEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFN 94  
 Db 111 EARKKPLPLGCHDIIIPVLPMFSKAVEHGK-----PSFTWGPTRVWISDESIR 162  
 QY 95 HVMKEA---YDYPKPGMAARVLIATGDGVVTAEGEAHKHRRIMIPSLSAQAVKSMVPI 151  
 Db 163 EVMSNKGHYGKPKPTRLGKLL-----ASGVVSVEGEKWKAKHRRILNPAFHHEKIKRMLPV 218  
 QY 152 FLEKGMELVDKMDAEDAAEKMAVGESAGEKATKATLETEGVDVVDKDWVGRATLDVWALAGF- 210  
 Db 219 FSNCTETWTRW-----ENSMI-EGMSE-----VDVWPEFQNLTDGVISKTAFG 262  
 QY 211 -DYKSD-----SLQNKTNELYVAFGLTDGFAPTLDSFKAIMWDFVPYFRMKRRHEIPT 265  
 Db 263 SSYEGRRIIFQLOAESAEIRII-----QAFRTI-----FIPGYWFLPTKNNRRLR 306  
 QY 266 QGLAVSRVGIEMLEQKQAVLGSASDAQVDKQVGRDILSLVRANI-AANLPESOKL 324  
 Db 307 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLDGLLVESNMRESNGKAEI 357  
 QY 325 SDEVLAQISNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREEIC-QIDTDMPTLDELN 383  
 Db 358 TTDEIIIECKLFYFAGMETTSVLLTTLVLSMHPWQERAREEVLHFGRTTPDYDSLS 417  
 QY 384 ALPYLEAFVKESLRDPPSPYANRECKLDEDFIPLAEPVIGRDSGVINEVRITKGTWML 443  
 Db 418 RLKIVTMLIYEVRLPVVFLTRTYKWEML-----GGIKYPAEVT-----LML 462  
 QY 444 PLFNINRSKFIYGEDAEERFRERWLEDVDTSLNSIEAPYGHQASFTIS---GPRACFGWR 500  
 Db 463 PILFIHDDPDWKGADGEFNGRPFADGISNATK-----YQTSFPFGWGPICIGQNF 515  
 QY 501 AVAEMKAFVTLRRVQPEPIIISHPEYH---ITLIISRPRIVGREKEGYQWRLO 552  
 Db 516 ALLEAKMAICTILQRFSE---LSPSYTHAPFTVITLHP-----OHGAQIKLK 560

RESULT 12  
 US-09-992-901-2  
 ; Sequence 2, Application US/09992901  
 ; Patent No. US20020073446A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Neff, Michael M.

```

; APPLICANT: Chory, Joanne
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
; FILE REFERENCE: SALKINS.024DV1
; CURRENT APPLICATION NUMBER: US/09/992,901
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/527,073
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170,931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172,832
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-992-901-2

```

```

Query Match      12.9%; Score 367; DB 9; Length 520;
Best Local Similarity 22.2%; Pred. No. 4.6e-24;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

QY 4 LVLLTGALGAAFAFWASTAFPSLYLAPRR-----SSLYNLOQPNHNTVFTGNFLDILS--- 56
DB 13 VLVLSVILSVIVGMSL-----LWRRPKIEEHFSGKIGPPY-HFFGNVVELVGMWL 67
QY 57 -----ARTGEEHAKYREYXGTLRFAGIAGAPVINST--DPKVFNVHMKEA 100
DB 68 KASSHEMPFESHNIPRLVSLFVHHWKIVGATFL---VWFGTFLRVADPDLIREIFSKS 124
QY 101 YDYPKGAARVLRATGCVVTAEGEAKHEHRIMPSLSAQAVKSMVPIFELEKMEVLV 160
DB 125 EFYEK-NEAHLPLVQLGEGDGLSLKGEKWAHRKILISFTHEMNLKLVFVVLKSVTDWV 183
QY 161 DKQMDAAEKDMVAGESAGEKATRLTEGVDVVDXDWVGRATLDVMALAGFDYKSDSLQNK 220
DB 184 DKMSDKLSEN-----GEVY-----VDVVEWFQILTEDVISRTAFGSYE----- 222
QY 221 TNELTVAFVGLTGDGAP-----TLDSFKAIMDFVPPYRTMKRHEIPLTQGLAV 270
DB 223 -----DGRAVFRLOAQOMLLCAEAFQVY---FIPGYRFP-----TRGNLK 260
QY 271 SRRVG-----IELMEOKQAVLGSASDOAVDKKQVQGRDILSLVRAIAANLPESQK 323
DB 261 SKLDKEIKRKSLLKLIERRQNAIDGEGEC---KEPAKDLGLMIQA-----KN 308
QY 324 LSDEVLQAQINLLFAGYETSTVLVTFHRLSEKAVQDKLREICQI---DTDMPTLDE 381
DB 309 VTVQDIVEECKSFFFAKQTTSNLLTWTLLSMHPWQAKARDEVLRVCGSRDVTCKH 368
QY 382 LNALPYLEAFVKESLRLLDPPSPYANRECLKDEDPIPLAEPVIGRDSGVINEVITKGTWV 441
DB 369 VVKLTSLMILNESRLPPIVATIRAKSDVK-----LGGYKIPGCTEL 413
QY 442 MLPLFNINRSKFYIGEDAEFRPERWLEDVTDLSNLSIAPYGHQASFSIS---GPRACFGW 498
DB 414 LIPITAVHDDQAIWGDVNEFNPARFADGVPRAAK-----HPVGIPFGLGVRTCIGQ 466
QY 499 RFAVEMKAFVTLRRVQFEPFIISHPEYEHITLII 534
DB 467 NLAILQAKLTAVMTICRFTFH---LAPTYQHAPTVL 499

```

```

RESULT 13
US-10-425-114-38827
; Sequence 38827, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

```

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38827
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700045236_FLI.pap
US-10-425-114-38827

```

```

Query Match      12.8%; Score 365.5; DB 12; Length 662;
Best Local Similarity 22.9%; Pred. No. 9.1e-24;
Matches 123; Conservative 96; Mismatches 193; Indels 125; Gaps 16;

QY 70 YGSLRFAGTAGAPVINSTDPKVFNVHMKEAIDYKPGMAARVLRATGCVVTAEGEAH 129
DB 188 YGGIFRL-NFGPKSFLIVSDPAIAGHILRENSKAYSKGILAEILFEVMTGTLIPADGEIM 246
QY 130 KRURRIMPISLSAQAVKSMVPIFELEKMEVLVDMEDAAEKDMVAGESAGEKATRLETE 189
DB 247 RVRRRAIVPALHOKYVTAMIGLFGEASQRLCEKL-----DKAA--VDGE 288
QY 190 GVDVKWVGRTLDVMALAGFDYKSDSLQNKINELYVAVFVGLTDFAPTLDSFKAIMWDF 249
DB 289 DMEMESLFSRLTLDVTKAVFNDFDSLDYDNGIVEAVVTVIRE--AEMRSTSPITW-- 344
QY 250 VVPYRTMKRHEIPLTQGLAVSRVVGIELMEOKQAVLGSASDQ-----AVDKKDVQG 302
DB 345 -----EPIWKDISPROKTYNEALK-----LINSTLDELIAICKRLVEQEDLQF 388
QY 303 RD-----ILSLVRAIAANLPESOKLSDEEVLQAQINLLFAGYETSTVLTWMPH 353
DB 389 HEYMNEQDPSILHFL-----LASGDDVSKQRLDDMTMLTAGHETSAVLTWTFY 440
QY 354 RLSEDKAVQDKLREICQIDTD-MPTLDELNALPYLEAFVKESLRLLDPPSPYANRECLKD 412
DB 441 LLSKPKVMKLODEADSVLGDGLPIEDVKLKYTRVINESLRLYPQPPVILRESLED 500
QY 413 EDFIPLAEPVIGRDSGVINEVITKGTWMLPLFNINRSKFYIGEDAEFRPERWLEDVT 472
DB 501 D-----ILGGYPIGRGEDIPISVWNLHHC9-KHWDDBAEVFNPERWPLDGP 544
QY 473 DSLNSIEAPYGHQASFSISGPRACFGWRFAVEMKAFVTLRRVQFE----- 519
DB 545 NP-NEINQNSY-LPFGGGRKCVGDMFATFEVTVATMLVKRFDQFQMAPGAPPVDMWTG 602
QY 520 -----PIISHPEYEHIT-----LIISRPIVGREKEG 546
DB 603 ATIHTTEGLKMTVTRRTPPVIPNLENKIITDSQESTLSAPSVMVSAASVASEGQ3 659

```

```

RESULT 14
US-10-114-270-60
; Sequence 60, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spyttek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.

```

APPLICANT: Li, Li  
APPLICANT: Vernet, Corine  
APPLICANT: Zerhusen, Bryan D.  
APPLICANT: Gorman, Linda  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Pena, Carol E.A.  
APPLICANT: Smithson, Glennda  
APPLICANT: Burgess, Catherine E.  
APPLICANT: Gerlach, Valerie  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Shimkets, Richard A.  
APPLICANT: Gangolli, Bsha A.  
APPLICANT: Taupier Jr., Raymond J.  
APPLICANT: Casman, Stacie J.  
APPLICANT: Ji, Weizhen  
APPLICANT: Anderson, David W.  
APPLICANT: Liette, Mario W.  
APPLICANT: Rastelli, Luca  
APPLICANT: Edinger, Shlomit R.  
APPLICANT: Stone, David J.  
APPLICANT: MacDougall, John R.  
APPLICANT: Rothenberg, Mark E.  
TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-322C  
CURRENT APPLICATION NUMBER: US/10/114,270  
CURRENT FILING DATE: 2002-11-27  
PRIOR FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: 60/281,086  
PRIOR FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: 60/281,136  
PRIOR FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: 60/281,863  
PRIOR FILING DATE: 2001-04-05  
PRIOR APPLICATION NUMBER: 60/281,906  
PRIOR FILING DATE: 2001-04-05  
PRIOR APPLICATION NUMBER: 60/282,020  
PRIOR FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: 60/282,930  
PRIOR FILING DATE: 2001-04-10  
PRIOR APPLICATION NUMBER: 60/282,934  
PRIOR FILING DATE: 2001-04-10  
PRIOR APPLICATION NUMBER: 60/283,512  
PRIOR FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/283,710  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: 60/284,234  
PRIOR FILING DATE: 2001-04-17  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 470  
SEQ ID NO 60  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-114-270-60

Query Match 12.8%; Score 365; DB 12; Length 520;  
Best Local Similarity 24.6%; Pred. No. 6.9e-24;  
Matches 143; Conservative 104; Mismatches 221; Indels 114; Gaps 23;

QY 2 FILVLLTGALGLAA--FSWASIAFFSLYLAPRSSLYNQ---GNHTYFTGNFLDILS 56  
DB 19 WLLLLVGLASWLLARILAWT---YTFY-----DNCCRLQCFPPQPKQNFVWGLGLVTP 69

QY 57 ARTGEH-----AKREKY-----GSTLRPAGIAGAPVLNSTDPKFNHVMKEAYDPKPG 107  
DB 70 TEEGKTLTQVAVYPOGFKVWLGPIIPFVILCHPTIRSI-----TNASAAIVPKDN 122

QY 108 MAARVLRITATGDGVVTAEGEAHKHRRIMTPSLSAQAVKSMVPIFLEKGMELVDKQMEDA 167  
DB 123 LFYGFLEKWLGDGLLSGDKWSHRRMLTPAFHFNILKSYITIFNKSANIMLDKQWHLA 182

QY 168 AEKDMVAGESAGEKATFLETEGVVDKWDVGRATLDVMALAGFDYKSDSQNKTNELYVA 227  
DB 183 SE-----GSSR-----LDMFEHISLMTLSQKCVSFES-NCQKPSF-YIA 223

QY 228 FVGLTDGFAPLDSFKAIMKDFVPYF-----RTMKRRHEIPLTQGLAVSRR-----V 274  
DB 224 AILELSAFVEKRNQOILHIDFLYLLTPDQQRHRACLVHD--FTDAVIGERRRLPSQ 281

QY 275 GIELMEQKQAVLGSASDAQVDKDKVQGRDILSLAVRANIANLPSQKLSDEEVLAQIS 334  
DB 282 GVDDBFLQAK-----AKSKTLDFDV-----LLSKD-----EDGKLSDEEDIRAEAD 323

QY 335 NLLFAGYETSSTVLTWTFHRLSEDKAVQDKLREEICQIDTDMPTL-----DELNALPYLEA 390  
DB 324 TFMPEGHDTTASGLSWLYHLAKHPEYQERCRQVEQLLAKDRPIEIEWDDLAQLFLTM 383

QY 391 FVKESLRLDPPSPYANRECKLDEDFIPLAEPVIGRGSVINEVRIITKGTWMLPLFNINR 450  
DB 384 CIKESLRLLHPVPVISRGCTQD-----FVLPGRV-----IPKGIICLSVFGTHH 429

QY 451 SKFIYGEDAEFRPRERWLEDVTDLSNLSIEAPYGHQASFISSPRACFGWRFAVAKFLF 510  
DB 430 NPAVW-PDEYVDPRF-----DPENIKERSPLAFIPFSAGPRNCIGQTFAMAEMKVLA 483

QY 511 VTLRRVQPEPIISHPEVEHITLIISRPRIVYREKEGYQMLQ 552  
DB 484 LTLIRFVLPDHTPR-----RKPVLVRAEGGLWLRVE 517

RESULT 15  
US-10-424-599-199559  
; Sequence 199559, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 199559  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MKT3847\_22226C.1.pep  
US-10-424-599-199559

Query Match 12.4%; Score 353; DB 12; Length 524;  
Best Local Similarity 24.6%; Pred. No. 8.5e-23;  
Matches 145; Conservative 89; Mismatches 215; Indels 140; Gaps 25;

QY 1 MFILVLLTGALGLAAFSWASIAFFSLYLAPRSSL-----YNLQGNHTNYFTGNFLDILS 56  
DB 13 VFVIVI-----LALTSAWRVNLNWLIRPKLERLLREQLQG-NPYRILVGLKEIVK 65

QY 57 ARTGEHAK-----YREKYGSTLR-----FAGIAGAPVLNSTDPKFNHVMKE 99  
DB 66 LQM-EARSKPMNLSDHIVPRVFAHLHQSVLKHGKNSFTWFGPKPRVTLTDPDLINKLNK 124

QY 100 AYDYPKGAARVLRITATGDGVVTAEGEAHKHRRIMTPSLSAQAVKSMVPIFLEKGMEL 159  
DB 125 ISDFRPEANPLAKLAT--GLVNYDGEKNKRLINLPAFSELEKLMFLIPFKSCNDL 182

QY 160 VDKWMDAAEKDMVAGESAGEKATFLETEGVVDKWDVGRATLDVMALAGFDYKSD---- 215  
DB 183 IIKW-----EGMLSVDGSC-----NDVAFPLQNLASDVIARTAFSSFEQGR 226

QY 216 --SLQNKTNELYVAVGLTGDGFAPLDSFKAIM-----WDFVYFRTMKRRHEIPLTQ 267  
DB 227 IFQLQKELABL-----TMKVINKYIPGWRFPV-TATNRNKEI----- 264

```

QY 268 LAVSRVGIEMEQKQAVLGSASDAQVDKXQVQGRDILSLVRA-----NIA 315
Db 265 ---DRIKASLTDMIK-----REKAPKTGEATRDLLGILLESNHKEIQEHRNNVNG 315
QY 316 ANLPESQKLSDEEVLQISNLLFAGYETSTVLTWTFHRLSEDKAVQDKLREIQQI-DT 374
Db 316 MNL-----NDVIECKLYFAGQETTSVLLVWTVLLSRYDPWQSRABEVLQVFGK 367
QY 375 DMPTLDELNALPYLE-AFVKESLRDPPSPYANRECKLDEDFIPLABPVIGRGGSVINEV 433
Db 368 QAPNFDGLSHLKIVTMIELKVLRLYPPAVGLNEN-----VDRDMKLG-N-L 412
QY 434 RITKGTWVWMLPLFNINRSKFIYGEDAEFFRPERWLEDVTDLSNLSIA--PYGHQASPTSG 491
Db 413 SLPAGVQVSLPTTVPVPHDELWGDVNEFFKPERSEGVKATNGRVSPFFFGW-----G 466
QY 492 PRACFGWRPVAEMKAFLEVTIRRVQPEPIISHPEYEH--ITLIISRPR 538
Db 467 PRICIGQNFSLLEAKVALSTILOHFSFE---LSPAYAHAFVTVFTLQPK 512

```

Search completed: April 2, 2004, 14:10:36  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17.5 Seconds  
(without alignments)  
1643.181 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852  
Sequence: 1 MFILVLLTGALGAAFSWAS.....RIVGREKEGVQKRLQKVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTCUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	4	US-09-518-386B-1
2	2852	100.0	557	4	US-09-518-386B-3
3	396.5	13.9	503	4	US-09-144-367-2
4	370	13.0	504	1	US-08-457-274A-25
5	370	13.0	504	5	PCT-US95-05758-25
6	367	12.9	520	4	US-09-527-073-2
7	345	12.1	524	4	US-09-976-594-533
8	338.5	11.9	503	4	US-09-583-447A-2
9	331	11.6	540	4	US-09-302-620B-99
10	330.5	11.6	507	1	US-08-457-274A-22
11	330.5	11.6	507	5	PCT-US95-05758-22
12	328	11.5	504	4	US-09-583-447A-4
13	328	11.5	540	4	US-09-302-620B-98
14	326.5	11.4	541	3	US-09-158-767-19
15	326.5	11.4	541	3	US-09-158-767-20
16	321	11.3	576	3	US-08-948-564-16
17	318.5	11.2	526	1	US-08-298-426-4
18	294.5	10.3	489	4	US-09-852-067-4
19	294	10.3	522	4	US-09-302-620B-97
20	289	10.1	522	4	US-09-302-620B-96
21	276	9.7	420	4	US-09-583-447A-6
22	276	9.7	467	4	US-09-126-420A-17
23	269.5	9.4	517	4	US-09-302-620B-100
24	269.5	9.4	517	4	US-09-911-781-32
25	269	9.4	507	1	US-08-457-274A-23
26	269	9.4	507	5	PCT-US95-05758-23
27	267.5	9.4	517	4	US-09-302-620B-101

28	263	9.2	523	4	US-09-302-620B-95	Sequence 95, Appl
29	259.5	9.1	510	3	US-08-948-564-4	Sequence 4, Appl
30	259	9.1	510	4	US-09-852-067-2	Sequence 2, Appl
31	251.5	8.8	508	4	US-09-126-420A-25	Sequence 25, Appl
32	246	8.6	512	4	US-09-302-620B-103	Sequence 103, Appl
33	243	8.5	512	4	US-09-302-620B-102	Sequence 102, Appl
34	241.5	8.5	476	1	US-08-333-075A-30	Sequence 30, Appl
35	236.5	8.3	498	1	US-08-457-274A-24	Sequence 24, Appl
36	236.5	8.3	498	5	PCT-US95-05758-24	Sequence 24, Appl
37	235.5	8.3	499	4	US-09-302-620B-104	Sequence 104, Appl
38	235.5	8.3	504	4	US-09-126-420A-18	Sequence 18, Appl
39	233.5	8.2	509	4	US-09-499-302A-8	Sequence 8, Appl
40	228	8.0	426	2	US-08-560-398-4	Sequence 4, Appl
41	227	8.0	504	4	US-09-499-302A-6	Sequence 6, Appl
42	226.5	7.9	500	3	US-09-292-768-68	Sequence 68, Appl
43	226.5	7.9	500	3	US-09-292-768-70	Sequence 70, Appl
44	225.5	7.9	500	3	US-09-292-768-4	Sequence 4, Appl
45	223.5	7.8	524	4	US-09-126-420A-24	Sequence 24, Appl

## ALIGNMENTS

RESULT 1  
US-09-518-386B-1  
; Sequence 1, Application US/09518386B  
; Patent No. 6365386  
; GENERAL INFORMATION:  
; APPLICANT: HOSHINO, Tatsuo  
; APPLICANT: OJIMA, Kazuyuki  
; APPLICANT: SETOGUCHI, Yutaka  
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
; CURRENT APPLICATION NUMBER: US/09/518.386B  
; CURRENT FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: EP 99104668.1  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EP 00101666.6  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 557  
; TYPE: PRT  
; ORGANISM: Phaffia rhodozyma  
; FEATURE:  
; NAME/KEY: TRANSIT  
; LOCATION: (1)..(26)  
US-09-518-386B-1

Query Match	100.0%;	Score 2852;	DB 4;	Length 557;
Best Local Similarity	100.0%;	Pred. No. 4.4e-278;		
Matches 557;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MFILVLLTGALGAAFSWASIAFESLYLAPRRSLYNLQGNHNTFTGNFLDLSARTG	60	
DB	1	MFILVLLTGALGAAFSWASIAFESLYLAPRRSLYNLQGNHNTFTGNFLDLSARTG	60	
QY	61	BEHAKYREKYGSTLRFAGIAGAPVLNSTDKPVNHWKEAYDPKPGMAARVLRTATGDG	120	
DB	61	BEHAKYREKYGSTLRFAGIAGAPVLNSTDKPVNHWKEAYDPKPGMAARVLRTATGDG	120	
QY	121	VVTAGEAHKRRHRIIMPISLQAQVKSMPVIFLEKGMELVDKMWDAAEKDMVAVGESAGE	180	
DB	121	VVTAGEAHKRRHRIIMPISLQAQVKSMPVIFLEKGMELVDKMWDAAEKDMVAVGESAGE	180	
QY	181	KKATLETEGVVDKDWGRATLDYMALAGDYKSDSLQNKTNELYVAFVGLTDGFAPTLD	240	
DB	181	KKATLETEGVVDKDWGRATLDYMALAGDYKSDSLQNKTNELYVAFVGLTDGFAPTLD	240	
QY	241	SFKAIWMDFPVYFRTMKRHEIPLTOGLAVSRVGVLEMEQKQAVLGSASQAVKQDV	300	
DB	241	SFKAIWMDFPVYFRTMKRHEIPLTOGLAVSRVGVLEMEQKQAVLGSASQAVKQDV	300	

QY 301 QGRDILSLLVRANAAANLPSOKLSDEEVLAAQISNLLFAGYETSSTVLTWMPHRLSEDKA 360  
 Db 301 QGRDILSLLVRANAAANLPSOKLSDEEVLAAQISNLLFAGYETSSTVLTWMPHRLSEDKA 360  
 QY 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420  
 Db 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420  
 QY 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480  
 Db 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480  
 QY 481 PYGHOASPIGPRACFCGWRFAVAMKAFVLTLLRRVQFEPPIISHPEYEHITLIISRPRIV 540  
 Db 481 PYGHOASPIGPRACFCGWRFAVAMKAFVLTLLRRVQFEPPIISHPEYEHITLIISRPRIV 540  
 QY 541 GREKEGYQMRLQVXPVE 557  
 Db 541 GREKEGYQMRLQVXPVE 557

RESULT 2

US-09-518-386B-3  
 ; Sequence 3, Application US/09518386B  
 ; Patent No. 6365386  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOSHINO, Tatsuo  
 ; APPLICANT: OJIMA, Kazuyuki  
 ; APPLICANT: SETOGUCHI, Yutaka  
 ; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
 ; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
 ; CURRENT APPLICATION NUMBER: US/09/518,386B  
 ; CURRENT FILING DATE: 2000-03-03  
 ; PRIOR APPLICATION NUMBER: EP 99104669.1  
 ; PRIOR FILING DATE: 1999-03-09  
 ; PRIOR APPLICATION NUMBER: EP 00101666.6  
 ; PRIOR FILING DATE: 2000-02-01  
 ; NUMBER OF SEQ ID NOS: 32  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 557  
 ; TYPE: PRT  
 ; ORGANISM: Phaffia rhodozyma  
 US-09-518-386B-3

Query Match 100.0%; Score 2852; DB 4; Length 557;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-278;  
 Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MPILVLTGALGLAFAWASIAFSLYLAPRRSSLYNQPNHTNYFTGNFLDILSARTG 60  
 Db 1 MPILVLTGALGLAFAWASIAFSLYLAPRRSSLYNQPNHTNYFTGNFLDILSARTG 60  
 QY 61 BEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHVMKEAYDYPKPGMAARVLRATG 120  
 Db 61 BEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHVMKEAYDYPKPGMAARVLRATG 120  
 QY 121 VVTAEGEAHKKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAEDAKOMAVGESAGE 180  
 Db 121 VVTAEGEAHKKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAEDAKOMAVGESAGE 180  
 QY 181 KKAATELETGVVDKDWGRATLDVMAAGFDYKSDLSQNKTNELVYAFVGLTGDGAPTLD 240  
 Db 181 KKAATELETGVVDKDWGRATLDVMAAGFDYKSDLSQNKTNELVYAFVGLTGDGAPTLD 240  
 QY 241 SFKALMWDVFPYFTMTKRHEIPIUTQGLAVSRVGVIELMEQKQAVLGASDAQVDDKDV 300  
 Db 241 SFKALMWDVFPYFTMTKRHEIPIUTQGLAVSRVGVIELMEQKQAVLGASDAQVDDKDV 300  
 QY 301 QGRDILSLLVRANAAANLPSOKLSDEEVLAAQISNLLFAGYETSSTVLTWMPHRLSEDKA 360  
 Db 301 QGRDILSLLVRANAAANLPSOKLSDEEVLAAQISNLLFAGYETSSTVLTWMPHRLSEDKA 360

QY 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420  
 Db 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420  
 QY 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480  
 Db 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480  
 QY 481 PYGHOASPIGPRACFCGWRFAVAMKAFVLTLLRRVQFEPPIISHPEYEHITLIISRPRIV 540  
 Db 481 PYGHOASPIGPRACFCGWRFAVAMKAFVLTLLRRVQFEPPIISHPEYEHITLIISRPRIV 540  
 QY 541 GREKEGYQMRLQVXPVE 557  
 Db 541 GREKEGYQMRLQVXPVE 557

RESULT 3

US-09-144-367-2  
 ; Sequence 2, Application US/09144367  
 ; Patent No. 6432639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lichter, Jay  
 ; APPLICANT: Guido, Marco  
 ; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4  
 ; FILE REFERENCE: SEQ-12P  
 ; CURRENT APPLICATION NUMBER: US/09/144,367  
 ; CURRENT FILING DATE: 1998-08-31  
 ; PRIOR APPLICATION NUMBER: 60/058,612  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 503  
 ; TYPE: PRT  
 ; ORGANISM: H. sapiens  
 US-09-144-367-2

Query Match 13.9%; Score 396.5; DB 4; Length 503;  
 Best Local Similarity 28.7%; Pred. No. 5.8e-31;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
 QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQG--PNHTNY--FTGNFLDILSARTG-----E 61  
 Db 7 LAMETWLLAVSLVLLYLYGTHSHGLFKLGIPTPLPFLGN---ILSYHKGFCMFDM 63  
 QY 62 BHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHVM--KEAYD--PKPGMAARVLRAT 117  
 Db 64 CHKYGKVGWF--YDG--QOPVLAITDDPMIKTVLVKCYSVFTNRRPFGVGFMSA-- 117  
 QY 118 GDGVVTAEGEAHKKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAEDAKOMAVGES 177  
 Db 118 ---ISIADEEWKRLRSLSPFTTSGLKEMVPIIAQYGDVLRNLRBA----- 164  
 QY 178 AGEKKALETE--BGVDKDWGRATLDVMAAGFDYKSDLSQNKTNELVYAFVGLTGDGA 236  
 Db 165 -----ETGKVTILKDVFGYSMDVITSFGVNDLSLNNPD----- 201  
 QY 237 PTLDSFKAIM--WDFV--PYFTMTK--RRHEIPTQGLAV---SRRVGIELMEQKQAVLGA 290  
 Db 202 PVENTKLLRDFDFLSPFTTSITVFPFLIPLEVLNLCVFPREVTNFKRSYKR-----M 256  
 QY 291 SDQAVDKDVGQROILSLLVRANAAANLPSOKLSDEEVLAAQISNLLFAGYETSSTVLTW 350  
 Db 257 KESRLEDQKHEVDFLOLMIDSSQNSKETESHKALSDELVAQSIIFIPAGYETTSVLSF 316  
 QY 351 MPRHUSEDKAVQDKLREIEICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANRE 408  
 Db 317 IMYELATHPDQVQKQLEIEIDAVLFNKAPPTYDTVLQMSYLDVMVNETLRLFLIAMRLERV 376  
 QY 409 CLKDDDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWL 468



Db 377 CKDVE-----INGMFIKGVVWVMPYSALHRDP-KYWTPEKFLPERFS 420  
 QY 469 EDVTSLSN-IEAPYGHQASISGPRACFGWRPVAEMKAFVTLRRVQFEP 520  
 Db 421 KKNKNIDPIYITPFG-----SGPRNCIGNRFPALNNMKLALRVLQNFSPFX 467

RESULT 4  
 US-08-457-274A-25  
 ; Sequence 25, Application US/08457274A  
 ; Patent No. 5734086

GENERAL INFORMATION:  
 APPLICANT: Scott, Jeffrey G.  
 APPLICANT: Tomita, Takashi  
 TITLE OF INVENTION: Cycochrome P450lpr Gene and Its Uses  
 NUMBER OF SEQUENCES: 29  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
 STREET: P.O. Box 1051, Clinton Square  
 CITY: Rochester  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 14603

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/457,274A  
 FILING DATE:

CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 30,727  
 REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 716-263-1304  
 TELEFAX: 716-263-1600  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 504 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Rat  
 STRAIN:  
 DEVELOPMENTAL STAGE: Adult  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT:  
 US-08-457-274A-25

Query Match 13.0%; Score 370; DB 1; Length 504;  
 Best Local Similarity 26.2%; Pred. No. 2.7e-28;  
 Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY 7 LTCALGLAFAFWASIAFFSLY-LAPRRSSLYNLOQ-----PNHNTYFTGNF-LDI 54  
 Db 3 LLSALTLETWLLAVLVLLVYGFGRTHGLFKQGIQPKPKPLPFGTVLNYWGLWKFDV 62  
 QY 55 LSARTGEHAKREKXGSLRAGAPVLNSTDPKVFNVHM-KEAY-----DYPKP 106  
 Db 63 -----ECHKYKGIWG---LFDG--QMLPFAITDTEMIRKVLVKCFSVFTNRDFGPV 111  
 QY 107 GMAARVLRIATGDGVVTAEGEAKHRRIMIPSLSAQAKVSWPIFLEKGMELVDKMD 166  
 Db 112 GI-----MGKAVSVADEEKRYALLSPPTTSGRLEKMPFIIBQYGDILVYLKQE 163  
 QY 167 AAEKDVAVGESAGEKATLET-EGVDYKDWVGRATLDVMAAGFDYKSDSLQN----- 219

Db 164 A-----ETGKPVTKKVFAGYSMDVITSTSGVNVDSLNNPKDPFV 204  
 QY 220 -KTNEI-----YVAFVGLTDGFAPTLDSFKAIW--DFVPYFRTMKRRHEIPL 264  
 Db 205 ETKKLLRDFDPDLFLSVLFPFLT-----PIYMLNICHFPKDSIEFFK----- 250  
 QY 265 TOGLAVSRRVGIELMEQKQAVLGSASDAQVDKXVQGRDILSLVLRANLANLPSQ-K 323  
 Db 251 -----KFYVRMKETRLDSVQKHRV-----DFLOLMNAHNDSDKESHTA 290  
 QY 324 LSDEEVLAQISNLLFAGYETSTVLTWTFRLSEDKAVQDKLREICQI--DTDMPFLDE 381  
 Db 291 LSDMEITAGSIIFAGYEPTSSLSFVLHSLATHPTOKKLOEIDRALPNKAPTYDT 350  
 QY 382 LNALPYLEAFVKESLRDPPSPYANRECLDKEDFIPLAEPVIGRDGVSINEVITKGTWV 441  
 Db 351 VMEMEYLDVNLNETLRLYPIGNRLRVRVCKDVE-----INGVFMFKGSVV 395  
 QY 442 MLPLFNINRSKFIYGEDAEERPERWLEDVDTSLNS-IEAPYGHQASISGPRACFGWRP 500  
 Db 396 MIPSYALHEDPOHWPE-PEERPERFSKENKGSIDPVYVLPFG-----NGPRNCIGMRP 448  
 QY 501 AVAEMKAFVTLRRVQFEP 520  
 Db 449 ALNNMKLALTQVLQNFQFP 468

RESULT 5  
 PCT-US95-05758-25  
 ; Sequence 25, Application PC/TUS9505758  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cornell Research Foundation, Inc.  
 ; TITLE OF INVENTION: Cycochrome P450lpr Gene and Its  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
 ; STREET: P.O. Box 1051, Clinton Square  
 ; CITY: Rochester  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 14603  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/05758  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Goldman, Michael L.  
 ; REGISTRATION NUMBER: 30,727  
 ; REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 716-263-1304  
 ; TELEFAX: 716-263-1600  
 ; INFORMATION FOR SEQ ID NO: 25:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 504 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Rat  
 ; STRAIN: Unknown  
 ; DEVELOPMENTAL STAGE: Adult  
 ; POSITION IN GENOME:





QY 172 MAVGESAGEKATLETETGVVDKDVGRATLDVMAAGFDYKSDSLQNTNELYVAVGL 231  
Db 192 F-----RHVRKHGQOTDIOELFFRLTVDSATEFLFGESASLRDSD-----VGL 236  
QY 232 TDGFAPLDSFKALMWDVFYFRITMK--RRRHEIPTQGLAVSRVVGIELMEQKQOAVLGS 289  
Db 237 T-----PITKQPEG-RGDFADAFVSYQYQAYRFLQOQWYILN--GAFF--RKSIATVHK 287  
QY 290 ASDQAVK-----KDVQGRDILSLVRANIANLPSOKLSDBEVLA-QISNLLFRAGE 342  
Db 288 FADHYQAKALELTDLQKQGVFYEL-----AKQTRDPKVLDRDQLNLILVAGRD 339  
QY 343 TSSTVLTWMEHRLSEDAVQDKLREBI-----CQIDTDMPTLDELNALPYLEAFVKS 395  
Db 340 TTAGLLSFVVELSRNPEVFAKLREVENRFGLGESARVEBISFESLAKSCEYLKAVINEA 399  
QY 396 LRLDPPPPYANRECLDKDEDFIPLAEPVIGEDGSVINEVRITKGTWMLPLFNINRSKIY 455  
Db 400 LRLTSPVPHNFRVATRNIT-LPRGG--GKDG--CSPVVRKGVQVMYVIGTHRDPSIY 453  
QY 456 GEDAEERFRPRWLEDVTDLSNIEAPYGHQASFIQPRACFGWRFAVAEMKAFVTLIR 515  
Db 454 GADADVFRERWPEPTRLKWAYVP-----FNGSPRICLQCPALTEAS--YVTVRL 504  
QY 516 VQ-----FRPIISHPEYHITLIS 535  
Db 505 LQFEGNLSLOPNAEYPPKQLNTLTL 530

## RESULT 10

US-08-457-274A-22  
; Sequence 22, Application US/08457274A  
; Patent No 5734086  
; GENERAL INFORMATION:  
; APPLICANT: Scott, Jeffrey G.  
; APPLICANT: Tomita, Takashi  
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/457, 274A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 716-263-1304  
; TELEFAX: 716-263-1600  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 507 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Musca domestica  
; STRAIN: Rutgers

; DEVELOPMENTAL STAGE: Adult  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: Chromosome 5  
US-08-457-274A-22  
Query Match 11.68; Score 330.5; DB 1; Length 507;  
Best Local Similarity 23.58; Pred. No 2.6e-24;  
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;  
QY 10 ALGAAAFSWASIAFFSLYLAPRSSLYNOGNH--TNYFTGNFLDILSA-RTGEEHAKY 66  
Db 10 ALGV-----LASLALYFVRNFGYKRGIPHEEPHLVMGNVGRSLRSKYHIGIADY 62  
QY 67 REKYGSTLRFAGI--AGAPVLNSTDPKFNHYV--KEAYDYPKPGMAARVLRATIDGVVT 123  
Db 63 YRKPKGSGPAGIFLGHKPAAVLVLDKELKRVLLKDFSNFANRGLYYNEKODPLTGLHM 122  
QY 124 AEGBAHRRHRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKOMAVGESAGEKA 183  
Db 123 VEGEKWSRLTKLSPTFTAGKMKMYNTVLEVGQRLLEVME- 164  
QY 184 TRLETEG-VDVKDWVGRATLDVMAAGFDYKSDSLQ- 226  
Db 165 -KLEVSSELDMDRLARFNTDVIQSVAFIECNLSRNPEDHFLANGRKXSIEVPRNALIM 223  
QY 227 AFVGLTDGFAPLDSFKALMWDVFYFRITMKRHEIPTQGLAVSRVVGIELMEQKQOAV 286  
Db 224 AFI-----DSFPEL-----SRKLGMRVLPEDVHQF 248  
QY 287 LGSASDAQVD---KKDVQGRDILSLVRANIANLPSOK---LSDBEVLAQISNLLFA 339  
Db 249 FMSSIKETVDYREKNNIRNDFLDVLDELK--NNPESISKGLGTENELAAQVFPVFLG 305  
QY 340 GYETSSTVLTWMEHRLSEDAVQDKLREBI-----CQIDTDMPTLDELNALPYLEAFVKS 395  
Db 306 GFTSSSTMGFALYELAQNOQLDRLEEBVNEVFPQFKEDNISYDALMNPYPYLDVNLNET 365  
QY 396 LRLDPPPPYANRECLDKDEDFIPLAEPVIGEDGSVINEVRITKGTWMLPLFNINRSKIY 455  
Db 366 LRKYPVGSALTROTLDY-----VVPHPKYV---LPKGLTVFPVLGIHYDPDELY 413  
QY 456 GEDAEERFRPRWLEDVTDLSNIEAPYGHQASFIQPRACFGWRFAVAEMKAFVTLIR 515  
Db 414 -PNPEEFDPERFSPVMVKQDSVD-----WLGFGDGPNCIGMRFGKMQSRGLGALVIRH 467  
QY 516 VQFE-----PIISHPE 526  
Db 468 FRFTVCSKRTDIPMOINPE 485

## RESULT 11

PCT-US95-05758-22  
; Sequence 22, Application PC/TUS9505758  
; GENERAL INFORMATION:  
; APPLICANT: Corneil Research Foundation, Inc.  
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05758  
; FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1304  
TELEFAX: 716-263-1500  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 507 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Musca domestica  
STRAIN: Rutgers  
DEVELOPMENTAL STAGE: Adult  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Chromosome 5  
PCT-US95-05758-22

Query Match 11.6%; Score 330.5; DB 5; Length 507;  
Best Local Similarity 23.5%; Pred. No. 2.6e-24;  
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;

QY 10 ALGLAASWASIAFSLYLAAPRSSLYNLOGPNH--TNYTGNFLDILSA-RTGEEHAKY 66  
DB 10 ALGV-----LASLALFYVRNFGYWKRRGIPHEEPFLHVMGVNGLRKYHIGELIADY 62

QY 67 REKYGSTLRFAGI--AGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATGQVVT 123  
DB 63 YRKFGSGPPAGIFLGHKPAAVLVDKELRKVLKOPSNFANRGLYNEKDDPLTGLHLM 122

QY 124 AGEAHKRRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAVGESAGEKKA 183  
DB 123 VEGEKWRLSLKLSFTTAGOKMYNTVLEVGORLLEVME-----KINELY 164

QY 184 TRLETEG--VDVKDWGRATLDVMAAGFDYKSDSLQN-----KINELY 226  
DB 165 -KLEVSSELMEDILARENTDVIGSAFIECNSLRNPHDRFLAMGRKSTIEVPRHNALIM 223

QY 227 AFVGLTDFGATLSFKAIMDFVYPRFKRRHEIPTQGLAVSRVGVIELMEQKQAV 286  
DB 224 AFI-----DSFPEL-----SRKLGMRVLPEVDHOF 248

QY 287 LGSASDAQVD--KKDVGRDILSLVRANIAANLPESQK-----LSDEEVLAQISNLLFA 339  
DB 249 FMSSTKEIVDYREKNIRNDFDLVDLK---NNPESISKLGITFNELAAQVFFVFLG 305

QY 340 GYETSSTVLTWVHRLSDEKAVQKLRREI-----CQIDTDMPTLDELNALPYLEAFVKES 395  
DB 306 GFETSSTMGFALYELAQNLQRLREEVNEVDFQKEDNISYDALMNIPIYDQVLNET 365

QY 396 LRLDPPSPYANRECKDEDFPLAEPVLGRDGSVINEVRITKGTMMVLPFNINRSKFY 455  
DB 366 LRKYVPGALTKQTLNDY-----VPHNPKV-----LPKGTILVPIVLGIHYDELY 413

QY 456 GDAEEFRPERWLEDVTPSLNSIEAPYGHQAFISGPRACFGWFAVEMKAFVTLRR 515  
DB 414 -PNPEEDFERFSPBWVQRDSVD-----WLGFGDGRNRCIGMRFQVQSRGLGLALVIRH 467

QY 516 VQFE-----PIISHPE 526  
DB 468 FRFTVCSRTDIPMQINPE 485

RESULT 12  
US-09-583-447A-4  
; Sequence 4, Application US/09583447A

Patent No. 6645745  
; GENERAL INFORMATION:  
; APPLICANT: WOJNOWSKI, Leszek  
; APPLICANT: GELLNER, Klaus  
; APPLICANT: EISEL, Regina  
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A  
; FILE REFERENCE: 310115.401  
; CURRENT APPLICATION NUMBER: US/09/583,447A  
; CURRENT FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-583-447A-4

Query Match 11.5%; Score 328; DB 4; Length 504;  
Best Local Similarity 27.1%; Pred. No. 4.6e-24;  
Matches 127; Conservative 75; Mismatches 185; Indels 82; Gaps 18;

QY 68 EKYGSTLRPAGI--AGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATGQVVT 124  
DB 66 EKYG---EMWGLYEGQQPMLVIMDEPMIKTVLKECYSVFTNQMLPGMGLK-SALSFA 121

QY 125 EGEAHRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAVGESAGEKKA 184  
DB 122 EDEWKRIRTLSPAFSTVKFKEMVPIISQCQDMLVRSURQAE----- 165

QY 185 RLTEGVDVMDWGRATLDVMAAGFDYKSDSLQNTNELYVAVFGLTDGFAPTLDSPKA 244  
DB 166 --NSKINLKDFGAYTMDVITGTLGVNLDLNNPQD-----PFLKNMKK 209

QY 245 IM-WDFVPRTRMKRRHEIPTQGLA-VSRRVGIELMEQKQAVLGSASDAQVDK--XDV 300  
DB 210 LLKLDFLDPPLL-----ISLFFPLTPPEALNIGLFPKOVTHFLKNSIERWKESRLKDK 264

QY 301 QGR--DILSLVRANIAANLPESQKLSDEEVLQISNLLFAGYETSSTVLTWVHRLSED 358  
DB 265 QKRVDFEQMIQSDNSKTSKSHKALSDLELVAQSIIFAAVDTTSTLTPIMVELATH 324

QY 359 KAVQDKLRREICQIDTDM-----TLDELNALPYLEAFVKESLRDLDPSPYANRECKDE 413  
DB 325 PDVQQLQEE---IDAVLPNKAPVTYDALVQMEYLDMMVNETLRLFPVVSRTVRCKKDI 381

QY 414 DFPLAEPVLGRDGSVINEVRITKGTMMVLPFNINRSKFYGEDAEFRPE-RWLEDVT 472  
DB 382 E-----INGVFIKGLAVMVPYIALHHDH-P-KYWTPEKFCPSRFSKXNK 425

QY 473 DSLNSIE-APYGHQAFISGPRACFGWFAVEMKAFVTLRRVQFEP 520  
DB 426 DSDLYRYIPFG-----AGPRNCIGMRFALTNIKLAIVIRALQNFSPKP 468

RESULT 13  
US-09-302-620B-98  
; Sequence 98, Application US/09302620B  
; Patent No. 6331420  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, C. Ron  
; APPLICANT: Craft, David L.  
; APPLICANT: Erick, Budley  
; APPLICANT: Eshoo, Mark  
; APPLICANT: Madduri, Krishna M.  
; APPLICANT: Cornett, Cathy A.  
; APPLICANT: Brenner, Alfred A.  
; APPLICANT: Tang, Maria  
; APPLICANT: Loper, John C.  
; APPLICANT: Gleeson, Martin  
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450  
; TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA  
; TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS

```
; TITLE OF INVENTION: RELATING THERETO
; FILE REFERENCE: 1010-16 seq
; CURRENT APPLICATION NUMBER: US/09/302.620B
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 98
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-09-302-620B-98

Query Match      11.5%; Score 328; DB 4; Length 540;
Best Local Similarity 23.9%; Pred. No. 5.1e-24;
Matches 133; Conservative 93; Mismatches 197; Indels 134; Gaps 22;

QY 42 NHTNYF-----TGNFLDILSARTGEE--HAK-----YREKYG 71
DB 25 NTRWYFPLVLLSLNLSLHLYLERRPHAKPLGNFVDPTFGIATPLLIIYLSKG 84
QY 72 STLREA-----GIAGAPVNSTDPKVFNVHMKEAYDYPKPGNAAR 111
DB 85 TWKPFAGLWNNKXIVRDPKXKTTGLRIVGLPLIETMDPENIKAVLATQFNDFSLGTRHD 144
QY 112 VLRIATGCGVTAEGEAHKHRRIMPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKD 171
DB 145 FLYSLGDDGIFLTDGAGWKHSRTMLRQFAREQVSHV-----KLEPHVQVF 191
QY 172 NAVGESAGEKKATRLTEGVVDKWDVGRATLDVMAAGFDYKSDSLQNKTNELVYAFVGL 231
DB 192 F-----KHVRKRGQTFDIQELFFRLTVDSATEFLFGESASLRDES-----IGL 236
QY 232 TCGFAPTLDSFKAIMWDFVPPRTMK--RRHEIPLTQGLAVSRVGTIELMEQKQAVLGS 289
DB 237 T-----PTTKDFDG--RDFADAFNYQTYQAVRFLQQMYILN--GSEF--RKSIAVHK 287
QY 290 ASDQAVDK-----KDVQGRDILSLVRANTAAANLPESQKLSDEVLAA-QISNLLFAGYE 342
DB 288 FADHYVQKALELTDQDQGVVFLYEL-----AKQTRDPKVLDRQLNLVAGRD 339
QY 343 TSTVLTWPHRLSEDKAVQDKLREI-----CQIDTMDPTLDLNAFLYLEAFVKES 395
DB 340 TTAGLSFFVFLSRNPVFAKLREVENRPFGLGEARVEISFESLKSCEYLKAVINET 399
QY 396 LRLDPPSPYANRECLKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFYI 455
DB 400 LRLYPSVPHNFVATENTT-LRGG--GEDG--YSPIVVKQGVWYTVIATHRDPSIY 453
QY 456 GEDAEFRPRMLEDVTDLSINIEAPYGHQASFGSPRACFGWRFAVEMKAFVTLRLR 515
DB 454 GADADVFRPRMFEPEPTRKLGWAYVP-----FNGGPRICLQQQFALTEAS---YVTVRL 504
QY 516 VQPEPIISHPEYEHITL 532
DB 505 LQ-----EFAHLSM 513

RESULT 14
US-09-158-767-19
; Sequence 19, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; EARLIER FILING DATE: 1998-09-23
; EARLIER FILING DATE: 1997-09-24
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 98
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Candida tropicalis

Query Match      11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.6%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

QY 78 GIA--GAPVNSTDPKVFNVHMKEAYD--YKPGMAARVLRATGCGVTVTAEGEAHKHRR 134
DB 73 GVAARGGLVTVTCDPRLNLEHLVKARFDNYPKGFPMGVRFDLGLDGFISDGDGTWLAQSK 132
QY 135 IMIPLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDQAVGSAGEKKATRLTEGVVDVK 194
DB 133 TAALFETRLTIRTMSRWSRSIH--GRLLPILA--DAAKGKAQ-----VDLQ 176
QY 195 DWGRATLDVMAAGFDYKSDSLQNKTNELVYAFVGLTDGPAFLDSEFKAIMWDFVY 252
DB 177 DLLLRLTFDNLICGLAFGKDPETLAQGLPENEFAFDRATEA---TLNRF-----IFPE 227
QY 253 F-----RTMKERHEIPLTQGLA--VSRVGTIELMEQKQAVLGSASQAVDKDVQGRDIL 306
DB 228 FLWCKKWLGLMGMTTLLTSSMAHVDQVLAIVIKRLELAAGKCDTAATHD---DLL 283
QY 307 SLLVRANIAANLPESQKLSDEVLAAQISNLLFAGYETSTVLTWPHRLSEDKAVQDKUR 366
DB 284 SRFMKRG-----SVSDLSLQHVAFILAGRODTSSVALSWPFWLVSHPAVERKIV 334
QY 367 BEICQI-----DTDMPTLDLNAFLYLEAFVKESLRLDPPSPYANRECLKDE 413
DB 335 RELCSVLAASRGADHPALMAEPFTFEELRLVYLKAASETLRLYPSVPEDSKHVVD-- 393
QY 414 DFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFYIGEDAEFRPRMLE-- 469
DB 394 DYLP-----DOTF-----VPAGSSVTYSIYAGRMKGWGEDCLEFRPERWLSADGT 440
QY 470 --DVTDSLNSIEAPYGHQASFGSPRACFGWRFAVEMKAFVTLRLRVQFEPIISHPEY 527
DB 441 KFEQHDSYKFY-----AFNAGFRVCLGKLAYLQMKNIAGSVLLRHLRTVAPGRHVE 492
QY 528 EHITLISRRPRIVGREKEGYQMRLOQVP 555
DB 493 QKMSLTLEM-----XGG--LRMEVSP 511

RESULT 15
US-09-158-767-20
; Sequence 20, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; EARLIER FILING DATE: 1998-09-23
; EARLIER FILING DATE: 1997-09-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Altered sequences
US-09-158-767-20

Query Match      11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.6%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

QY 78 GIA--GAPVLNSTDPKVFNHVKAEYD-YKPCMAARVLRIATGQVVTAEGEAHKRHR 134
Db 73 GVARRGLVTVTCDFRNLEHVKARFNYPKGPFWGVFDLLGDI FNSDGDWTWLAQEK 132
QY 135 IMPLSAQAVKSMVPIFLKGMELVDKXMDAAEKMDAVGSAAGEKKATRLTEGVVK 194
Db 133 TAALFEFTTRTLRTAMSRWVSRSIH--GRLLPILA--DAAKGKAQ-----VDLQ 176
QY 195 DWGCRATLDVNALAGPDYKSDSLQN--KTNELYVAFVGLTDGFAPLDSFKAIMWDFVY 252
Db 177 DLLRLTFDNLGAFGKDPETLAQGLPENEPAFAFDRATEA---TLNRF-----IPPE 227
QY 253 F-----RTMKRRHEIPLTOGLA-VSRVVGIELMEQKKQAVLGSASDAQVDKDKVQGRDIL 306
Db 228 FLWRCKWGLGMEITLTSMAHVQVYLAIVKRRKLELAAGNGKCDTAATHD----DILL 283
QY 307 SLLVRANIAANLPESQKLSDEVLQAISNLLPAGYETSSVLTMMFHRLSDEKAYODKLR 366
Db 284 SRFNRKG-----SYDSELOHVALNFILAGROTSVALSWFFWLSTHPAVERKIV 334
QY 367 BEICQI-----DTDMPTLDLNALPYLEAFVKESLRLDPPSPYANRECLKDE 413
Db 335 RELCSVLAASRGADHPALWLAEPFPEELDRLVYKAALETURLYPSVPEDSKHVAD- 393
QY 414 DFIPLAEPVIGRDSVINEVRITGTWMLPLFNINRSKFIYGEDAEERPERWLE---- 469
Db 394 DYLP-----DGTG-----VPAGSSVTYSIVSAGRMKGWGEDCLEFRPERWLSADGT 440
QY 470 -DVTDSLNSIEAPYGHQAFISGFACFGWRFAVAKMAFLFTLRRVQFEPISHPY 527
Db 441 KFEQDSYKVV-----AFNAGPRVCLGKDLAYLQMKNIAGSVLLRHRLTVAPGRHVE 492
QY 528 EHITLIISRPRIVGREKEGYQMRLOVXP 555
Db 493 QKMSLTLEW-----KGG--LRMEVRP 511
```

Search completed: April 2, 2004, 14:02:14  
Job time : 19.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 54 Seconds  
(without alignments)  
2914.425 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLTGALGLAAFWAS.....RIVGREKGYQMLRQKPV 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2852	100.0	557	3	AAB08713 Amino aci
2	398.5	14.0	503	6	ABU57260 Human chy
3	398.5	14.0	503	6	ABR82024 Human cyt
4	396.5	13.9	502	7	AD448378 Human pro
5	396.5	13.9	503	2	AAR72363 Human cyt
6	396.5	13.9	503	2	AAR81464 Human der
7	396.5	13.9	503	2	AAR93170 Human cyt
8	396.5	13.9	503	2	AAY05202 Human CYP
9	394	13.8	503	5	ABG68753 Cytochrom
10	392.5	13.8	527	5	ABBO8079 Maize cyt
11	385.5	13.5	503	5	AAB22852 Human cyt
12	385.5	13.5	535	5	ABG68747 Cytochrom
13	385	13.5	502	5	AAB26192 Human cyt
14	385	13.5	502	7	ADE57186 Human pro
15	385	13.5	502	7	ADE57190 Human pro
16	385	13.5	502	7	ADBE3319 Human pro
17	382.5	13.4	454	5	ABG68754 Cytochrom
18	373.5	13.1	454	6	ADA48312 Rice prot
19	370	13.0	504	7	ADE57188 Rat prote
20	370	13.0	504	7	ADE57194 Rat prote
21	370	13.0	504	7	ADE57194 Rat prote
22	367	12.9	520	3	AAB23917 Arabidops
23	365	12.8	520	6	ABU54571 Human NOV
24	363.5	12.7	512	5	AAU97096 Abscisic
25	362.5	12.7	465	3	AAG54205 Arabidops

26	362.5	12.7	475	3	AAG54204	AAG54204 Arabidops
27	362.5	12.7	523	3	AAG54203	AAG54203 Arabidops
28	360	12.6	512	2	AAV05898	AAV05898 Vicia sat
29	356.5	12.5	475	3	AAG24685	AAG24685 Arabidops
30	356.5	12.5	522	3	AAG24684	AAG24684 Arabidops
31	356.5	12.5	527	3	AAG24683	AAG24683 Arabidops
32	354	12.4	562	4	AAU30319	AAU30319 Novel hum
33	351	12.3	512	5	AAU97097	AAU97097 Abscisic
34	350.5	12.3	1115	6	AAO16061	AAO16061 Gibberell
35	350.5	12.3	1054	6	AAO16055	AAO16055 Bacillus
36	350.5	12.3	1054	6	AAO16056	AAO16056 Human dru
37	349	12.2	508	5	ABO7528	ABO7528 Human dru
38	348.5	12.2	511	7	ADE60554	ADE60554 Human pro
39	348.5	12.2	546	4	AAU14298	AAU14298 Human nov
40	347.5	12.2	524	4	AAU14534	AAU14534 Human nov
41	347.5	12.2	524	4	AAU14536	AAU14536 Human nov
42	347.5	12.2	524	4	AAU14535	AAU14535 Human nov
43	346.5	12.1	511	7	ADE60552	ADE60552 Rat Prote
44	346.5	12.1	524	3	RAY66706	RAY66706 Membrane-
45	346.5	12.1	524	4	AAB73677	AAB73677 Human oxi

ALIGNMENTS

RESULT 1

AAB08713	standard; protein; 557 AA.
ID	AAB08713
XX	XX
AC	AAB08713;
XX	XX
DT	12-SEP-2003 (revised)
DT	02-JAN-2001 (first entry)
XX	XX
DE	Amino acid sequence of an astaxanthin synthetase polypeptide.
XX	XX
KW	Astaxanthin synthetase; astaxanthin; beta-carotene; carotenogenic yeast;
KW	antioxidant; cancer; colouring reagent; farmed fish; salmon.
XX	XX
OS	Xanthophyllomyces dendrorhous.
XX	XX
PN	EF1035206-Al.
XX	XX
PD	13-SEP-2000.
XX	XX
PF	03-MAR-2000; 2000EP-00104430.
XX	XX
PR	09-MAR-1999; 99EP-00104668.
PR	01-FEB-2000; 2000EP-00101666.
XX	(HOFF ) HOFFMANN LA ROCHE & CO AG F.
PI	Hoshino T, Ojima K, Setoguchi Y;
XX	WPI; 2000-559874/52.
DR	N-PSDB; AAA64472, AAA64473.
XX	Novel polynucleotide encoding astaxanthin synthase useful for producing
PT	recombinant cells for producing astaxanthin from beta-carotene.
XX	Claim 2; Page 21-23; 46pp; English.
PS	The present sequence represents an astaxanthin synthetase polypeptide of
XX	Phaffia rhodomyza. The enzyme is involved in the last step of the
CC	astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.
CC	rhodomyza is a carotenogenic yeast strain. The astaxanthin synthetase
CC	polynucleotides and polypeptides are useful for producing astaxanthin.
CC	Astaxanthin is an antioxidant which may be used to protect living cells
CC	against diseases such as cancer. Astaxanthin is also used as a colouring
CC	reagent, e.g. in farmed fish like salmon to impart an orange-red
CC	coloration. (Updated on 12-SEP-2003 to standardise OS field)
XX	Sequence 557 AA;



Query Match	100.0%; Score 2852; DB 3; Length 557;
Best Local Similarity	100.0%; Pred. No. 4.2e-240;
Matches 557; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MFILVLLTGALGAAFPASWASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDILSARTG 60
DB	1 MFILVLLTGALGAAFPASWASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDILSARTG 60
QY	61 EBAKREKYGSTLRPAGIAGAPVNSTDPKVFNHVMEAYDYKPKGVAARVLIATGGD 120
DB	61 EBAKREKYGSTLRPAGIAGAPVNSTDPKVFNHVMEAYDYKPKGVAARVLIATGGD 120
QY	121 VVTAEAEAKHRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMEADAEDKMAVGESAGE 180
DB	121 VVTAEAEAKHRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMEADAEDKMAVGESAGE 180
QY	181 KKATRLTEGVDKQWVGSRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDFAPTLD 240
DB	181 KKATRLTEGVDKQWVGSRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDFAPTLD 240
QY	241 SFKAIMWDFVFPRTMKRHEIPLTQGLAVSRVGIEMKQKQAVLGSASDAQVDKDV 300
DB	241 SFKAIMWDFVFPRTMKRHEIPLTQGLAVSRVGIEMKQKQAVLGSASDAQVDKDV 300
QY	301 QGRDILSLVVRANIANLPESQKLSDEEVLQISNLLPAGYETSTVLTWVHRLSDEKA 360
DB	301 QGRDILSLVVRANIANLPESQKLSDEEVLQISNLLPAGYETSTVLTWVHRLSDEKA 360
QY	361 VQKLEEEICQIDTMDPTLDENALPYLEAFVKESLRDLPSPVANRECKDEDFIPLAE 420
DB	361 VQKLEEEICQIDTMDPTLDENALPYLEAFVKESLRDLPSPVANRECKDEDFIPLAE 420
QY	421 PVIGRSGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPERWLEDVTDLSNIEA 480
DB	421 PVIGRSGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPERWLEDVTDLSNIEA 480
QY	481 PYGHOASFISGPRACGFWFAEAKELFVTLRRVQPEPIISHPEYEHITLIISRPIV 540
DB	481 PYGHOASFISGPRACGFWFAEAKELFVTLRRVQPEPIISHPEYEHITLIISRPIV 540
QY	541 GREKEGYQMLQVKPVE 557
DB	541 GREKEGYQMLQVKPVE 557
RESULT 2	
ABU57260	
ID	ABU57260 standard; protein; 503 AA.
XX	ABU57260;
AC	
XX	25-APR-2003 (first entry)
DT	
XX	Human chytochrome P450 CYP3A4 protein.
DE	
XX	Human; enzyme; transgenic; drug metabolism; behaviour; mouse;
XX	pharmacokinetic assay; pharmacodynamic assay; toxicology; serum albumin;
KW	alpha-acidic glycoprotein; CYP; multidrug resistance protein; MRP;
KW	uridine diphosphoglucuronosyl transferase, UGT; cytochrome P450.
XX	
OS	Homo sapiens.
XX	
XX	WO200283897-A1.
FN	
XX	24-OCT-2002.
PD	
XX	18-APR-2002; 2002WO-AU0004485.
EF	
XX	18-APR-2001; 2001AU-00004467.
PR	
XX	(GENE-) GENE STREAM PTY LTD.
PA	
XX	

PI	Daly JM;
XX	
DR	WPI; 2003-093021/08.
DR	N-PSDB; ABX77172.
XX	
PT	New transgenic non-human animal expressing a foreign polypeptide associated with drug behavior and/or metabolism, useful for studying the behavior and/or metabolism of a drug in other animals.
PT	
XX	Disclosure; Page 299-301; 408pp; English.
PS	
XX	
CC	This invention relates to a transgenic non-human animal which may be used for assessing the behaviour and/or metabolism of a drug in another animal and which expresses a foreign polypeptide associated with drug behaviour and/or metabolism. The invention also comprises a nucleic acid construct for use in producing the above transgenic non-human animal and a method of assessing the metabolism and/or behavior of a drug in an animal of interest, comprising administering a test agent to the transgenic animal and conducting analytical tests to determine drug metabolism and/or behavior. The transgenic animal is useful in studying drug metabolism and/or behavior in other animals. The nucleic acid construct is useful in producing the above transgenic animal and the methods are used for producing, breeding and using transgenic animals for pharmacological (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological studies. Nucleic acid sequences used within the invention are serum albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins and (MRP's). The present sequence represents a protein sequence used to create a transgenic animal within the scope of the invention
CC	
XX	Sequence 503 AA;
XX	
QY	Query Match 14.0%; Score 398.5; DB 6; Length 503;
DB	Best Local Similarity 28.7%; Pred. No. 8.6e-26;
QY	Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
DB	
QY	13 LAAFSW--ASTAFSLYLAPRRS--SLYNLQO--PNHTNY--FTGNFLDILSARTG-----E 61
DB	7 LAMETWLLAVSLVLLVLYGTHSHGLFKLGIGPPTLPFLGN---ILSVHKGFCMFDM 63
QY	62 EBAKREKYGSTLRPAGIAGAPVNSTDPKVFNHVMEAYDY---PKGMAARVLIAT 117
DB	64 CHKYGKVMGF---YDG--QQPLAITDDPMIKTVLVKECYSVFTNRRPPGPGVFKRSA- 117
QY	118 GDGVVTAEGEAHRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMEADAEDKMAVGES 177
DB	118 ---TSIAEDEBKRLSLLSPTFTSGKLEKEMVPIAQYGDVLYNLRREA----- 164
QY	178 AGEKATLET-EGVDVKDWVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDGFA 236
DB	165 -----ETGKPVTLKDVFGAYSDMTITSTSGVNIIDSLNNPQD----- 201
QY	237 PTLDSFKAIM-WDFV-PYFRTMK-RRHEIPLTQGLAV---SRVGIEMKQKQAVLGS 290
DB	202 PFVENTKLLRFDLPDFLSITVFPFLIPILEVANICVFPREVNTFLRSVKR-----M 256
QY	291 SDQAYDKDQVGRDILSLVVRANIANLPESQKLSDEEVLQISNLLPAGYETSSVLTW 350
DB	257 XESRLDTQKRVDFLQMLIDSQNSKETESHKALSDLELVAQSIIFIFAGYETSSVLSF 316
QY	351 MFHRLSEDKAVQDKLREICOI--DTDMPTLDENALPYLEAFVKESLRDLPSPVANRE 408
DB	317 IMTELATHPDVQOKLQEEIDAVLPKAPPYDYTVLQMEYLDVMVNETLRFPLAMLERV 376
QY	409 CLKDEDF:PLAEPVIGRSGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPERWL 468
DB	377 CKKQVE-----INGMFIPKGVVVMIPSYALHROP-KYWTBPEKFLPERFS 420
QY	469 EDVTDLSNS-IEAPYGHQASFISGPRACGFWFAEAKELFVTLRRVQPEPI 520
DB	421 KKNKONIDFYITTPG-----SGFRNCIGMRFALMNKALIRVLQNFSEK 467



CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 502 AA;

Query Match 13.9%; Score 396.5; DB 7; Length 502;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
 QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61  
 Db 6 LAMETWLLAVSLVLLYLYGTHSHGLFKLGIPOGTPPLFLGN---ILSYHKGFCMPDME 62  
 QY 62 EHAQYREKYGSTLRAGAPVLNSTDPKFNHVM--KEAYDY---PKPGMAARVLRIAT 117  
 Db 63 CHKYGKVGWF---YDG--QQPVLAITDPMIKTVLVEKCVFTNRRPFGVGFMSA- 116  
 QY 118 GDGVVTAEGEAHRRHRRIMIPSLSAQAVKSMVPIFLEKGMELVDQMDEAAEKDMAYGES 177  
 Db 117 ---ISIAEDEEWKRLSLLSPTFTSGKLKEMVPIIAQYGDVILVRLNREA----- 163  
 QY 178 AGEKKAIRLET-EGVDVKQWVGTRATLVNALAGPDYKSDSLONKTNELYAVFGLTDGFA 236  
 Db 164 -----ETGKPVTLKDVFGAYSDMTVITSTSGVNIIDSLNNPD----- 200  
 QY 237 PTLDSEKAIM-WDFV-PYFRTWK-RRHEIFLTQGLAV---SRRVGIELMEQKQAVLGS 290  
 Db 201 PFVENTKLLRFDLDFDPLFISITVFPFLIPILEVNICVFPREVTNLRKSVK- ---M 255  
 QY 291 SDQAVDKQVQGRDILSLVLRANIANLPESQKLSDEEVLQAISNLLFAGYETSSVLTW 350  
 Db 256 KESRLDQTKHRVDFLQMLDMSQSKETESHKALSDLELVAQSIIFIFAGYETSSVLSF 315  
 QY 351 MFHRLSDEKAVQKLRREEICQI--DTPMTLDELNALPYLEAFVXESLRLLDPPSPYANRE 408  
 Db 316 IMVELATHPDVQKQLEEDDAVLNPKAPPTVDTVLQMEYLDWMVNETLRLPIAMRLERV 375  
 QY 409 CLKDEDFIPLAEPVIGRDSVINEVRIITGTWVMPLFNINRSKIYGEDAEERFPERWL 468  
 Db 376 CKKDVE-----INGMFIKGVVMVPIPSYALHRDP-KYWTPEKEFLPERFS 419  
 QY 469 EDVTDLSNS-IEAPYGHQASFISGPRACFGWFAVAEMKAFVTLRLRVQEPF 520  
 Db 420 KKKONIDPIYIYTFPG-----SGPRNCIGMRFALMMKMLALIRVLQNFSEKFP 466

RESULT 5  
 AAR72363  
 ID AAR72363 standard; protein; 503 AA.  
 XX  
 XX AAR72363;  
 AC  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 14-NOV-1995 (first entry)  
 XX  
 XX Human cytochrome P450 molecular species 3A4 protein.  
 XX

KW Human cytochrome P450; amplification; PCR; primer; expression vector;  
 KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;  
 KW carcinogen; mutagen; liver metabolism.  
 XX Homo sapiens.  
 PN EP644267-A2.  
 XX  
 PD 22-MAR-1995.  
 XX  
 PF 20-JUL-1994; 94EP-00111298.  
 XX  
 PR 20-JUL-1993; 93JP-00201120.  
 PR 21-JUL-1993; 93JP-00180246.  
 PR 30-JUL-1993; 93JP-00208279.  
 XX  
 PA (SUMO ) SUMITOMO CHEM CO LTD.  
 XX  
 PI Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;  
 WIPI; 1995-116991/16.  
 DR N-PSDB; AAQ87717.  
 XX  
 PT Evaluation of safety of a chemical cpd. - using recombinant yeast  
 PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.  
 XX  
 PS Example; Page 31-33; 124pp; English.  
 XX  
 CC The amino acid sequence of the human cytochrome P450 species 3A4. The  
 CC cDNA was amplified by PCR using the primers AAQ87743-6. The product was  
 CC cloned into the yeast expression vectors pAAHNS or pAHRH to produce the  
 CC vectors p3A4 for the expression of the cytochrome P450 alone or p3A4R for  
 CC co-expression with the yeast NADPH-P450 reductase. The vectors are used  
 CC in a method for evaluating the safety of a chemical compound by reacting  
 CC the chemical compound with recombinantly produced human cytochrome P450  
 CC molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4,  
 CC or their auxiliary species and variants (AAQ87718-32), and yeast NADPH-  
 CC P450 reductase, either as a fused protein or in cell extracts, and  
 CC analysing the resulting metabolite to assess the safety of the chemical  
 CC compound. The method is useful for determining whether the chemical  
 CC compound, or its metabolite, will be converted into a carcinogenic or  
 CC mutagenic form through metabolism in the liver. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 XX  
 SQ Sequence 503 AA;  
 Query Match 13.9%; Score 396.5; DB 2; Length 503;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
 QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61  
 Db 7 LAMETWLLAVSLVLLYLYGTHSHGLFKLGIPOGTPPLFLGN---ILSYHKGFCMPDME 63  
 QY 62 EHAQYREKYGSTLRAGAPVLNSTDPKFNHVM--KEAYDY---PKPGMAARVLRIAT 117  
 Db 64 CHKYGKVGWF---YDG--QQPVLAITDPMIKTVLVEKCVFTNRRPFGVGFMSA- 117  
 QY 118 GDGVVTAEGEAHRRHRRIMIPSLSAQAVKSMVPIFLEKGMELVDQMDEAAEKDMAYGES 177  
 Db 118 ---ISIAEDEEWKRLSLLSPTFTSGKLKEMVPIIAQYGDVILVRLNREA----- 164  
 QY 178 AGEKKAIRLET-EGVDVKQWVGTRATLVNALAGPDYKSDSLONKTNELYAVFGLTDGFA 236  
 Db 165 -----ETGKPVTLKDVFGAYSDMTVITSTSGVNIIDSLNNPD----- 201  
 QY 237 PTLDSEKAIM-WDFV-PYFRTWK-RRHEIFLTQGLAV---SRRVGIELMEQKQAVLGS 290  
 Db 202 PFVENTKLLRFDLDFDPLFISITVFPFLIPILEVNICVFPREVTNLRKSVK- ---M 256  
 QY 291 SDQAVDKQVQGRDILSLVLRANIANLPESQKLSDEEVLQAISNLLFAGYETSSVLTW 350  
 Db 257 KESRLDQTKHRVDFLQMLDMSQSKETESHKALSDLELVAQSIIFIFAGYETSSVLSF 316

QY 351 MFHRLSEDKAVQDKLREIEICQI--DTDMPTLDELNALPYLEAFVKSLESLDPPSPYANRE 408  
Db 317 IMYELATHPDVQCKLOEEDAVLPNKAPPTDYTLQMEYLDVMVNETLRLFPPIAMRLERV 376  
QY 409 CLKDEDFIPLAEPVIGRDGVSINEVRITKGTVMVLPFLFNINRSKFIYGEDAEERPERWL 468  
Db 377 CKKQVE-----INGMFIPKGVVMIPISYALHRDP-KYWTEPEKFLPERFS 420  
QY 469 EDVTDLSNS-IEAPYGHQASPTSGPRACGWRFAVAEMKAFLEVTLRVQPEP 520  
Db 421 KKKNDIDPIYTPFG-----SGPRNCIGRFPALMMKALIRVLQNFSEKP 467

RESULT 6  
AAR81464  
ID AAR81464 standard; protein; 503 AA.  
AC AAR81464;  
XX  
DT 01-AUG-1996 (first entry)  
XX  
DE Human derived cytochrome P4503A4.  
XX  
KW Human derived cytochrome; P4503A4; commercial cDNA library; yeast;  
KW transfection; recombinant production; expression vector; mammal;  
KW immunisation; sensitisation; antibody; determination; detection;  
KW non-cross reactive.  
XX  
OS Homo sapiens.  
XX  
FN JP08027197-A.  
XX  
PD 30-JAN-1996.  
XX  
PF 13-JUL-1994; 94JP-00161552.  
XX  
PR 13-JUL-1994; 94JP-00161552.  
XX  
PA (SUMO) SUMITOMO CHEM CO LTD.  
XX  
DR WPI; 1996-136338/14.  
DR N-PSDB; AAT17399.  
XX  
PT Antibody recognising human derived cytochrome P4502A4 - allows specific  
PT detection of cytochrome P450 species in humans.  
XX  
PS Example 1; Page 10-12; 13pp; Japanese.  
XX  
CC The present sequence is the human derived cytochrome (HDC) P4503A4, which  
CC was obt'd. from a commercial cDNA library. Yeast were transfected with an  
CC expression vector contg. the HDC cDNA, cultured and then disrupted to  
CC give a microsomal fraction. The HDC was purified from the fraction, and  
CC used to immunise and sensitise a mammal. Blood was drawn from the mammal,  
CC and an anti-HDC antibody isolated. The antibody obt'd. recognises HDC  
CC P4503A4, partic. at a serum dilution rate of 1:10000, and is  
CC substantially without cross reaction to other HDC P450 spp  
XX  
SQ Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;  
Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAATSW--ASTAFSLYAPERS-SLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61  
Db 7 LAMETWLLAVSLVLLLYGTHSHGLPKKIGIGPPTPLPLGN---ILSVHKGFCNFDM 63  
QY 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHWM-KEAYDY---PKPGMAARVLRIAT 117  
Db 64 CHKXYGKWGF---YDG--QQPVLAITDPDMIKTVLVKCYSVFTNRPPFGPVGFMKSA- 117  
QY 118 GDGVVTAEGEAKHRRIMIFSLSAQAQKSVKVPFLEKGMELVDKWMEDAAEKDMAVGES 177

Db 118 ---ISTAEDENKRLSLLSPFTTSGKLKEMVPIIAQYGDVLRNLRREA----- 164  
QY 178 AGEKKATRLLET-EGVDVKDWGRALDVMALAGPDYKSDSLONKTNELYVAFVGLTDGFA 236  
Db 165 -----ETGKPTLKDVFAGYSMDVITSTSGVNI:DSLNNPOD----- 201  
QY 237 PTLDSFKAIM-WDFV-PYFRTMK-RRHEIPTLQGLAV---SRRVGIELMEQKQAVLGSA 290  
Db 202 PFVNTKLLARDFDLDPPLSITVPFPLIPILEVLNICVFPREVINFURKSVKR-----M 256  
QY 291 SDQAVKDKVQGRDILSLLVRANIANLPESOKLSDEVLQAGISNLLPAGYETSTVLTW 350  
Db 257 KESRLEDTQKRVDFLQMLMDSQNSKESHKALSDELVAQSIIFIPAGYETSTVLSLF 316  
QY 351 MFHRLSEDKAVQDKLREIEICQI--DTDMPTLDELNALPYLEAFVKSLESLDPPSPYANRE 408  
Db 317 IMYELATHPDVQCKLOEEDAVLPNKAPPTDYTLQMEYLDVMVNETLRLFPPIAMRLERV 376  
QY 409 CLKDEDFIPLAEPVIGRDGVSINEVRITKGTVMVLPFLFNINRSKFIYGEDAEERPERWL 468  
Db 377 CKKQVE-----INGMFIPKGVVMIPISYALHRDP-KYWTEPEKFLPERFS 420  
QY 469 EDVTDLSNS-IEAPYGHQASPTSGPRACGWRFAVAEMKAFLEVTLRVQPEP 520  
Db 421 KKKNDIDPIYTPFG-----SGPRNCIGRFPALMMKALIRVLQNFSEKP 467

RESULT 7  
AAR93170  
ID AAR93170 standard; protein; 503 AA.  
AC AAR93170;  
XX  
DT 11-OCT-1996 (first entry)  
XX  
DE Human cytochrome P450 molecular species 3A4 protein.  
XX  
KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic.  
XX  
OS Homo sapiens.  
XX  
FN JP08056635-A.  
XX  
PD 05-MAR-1996.  
XX  
PF 15-JUL-1994; 94JP-00164184.  
XX  
PR 20-JUL-1993; 93JP-00201120.  
PR 30-JUL-1993; 93JP-00208279.  
PR 17-JUN-1994; 94JP-00136053.  
XX  
PA (SUMO) SUMITOMO CHEM CO LTD.  
XX  
DR WPI; 1996-182311/19.  
DR N-PSDB; AAT28383.  
XX  
PT Novel method for the evaluation of the safety of a cpd. - using a human  
PT cytochrome P450 and yeast NADPH reductase to determine whether the  
PT analyte cpd. is detoxified or metabolised to a carcinogen.  
XX  
PS Example 1; Page 24-26; 74pp; Japanese.  
XX  
CC This is the amino acid sequence of the human cytochrome P450 molecular  
CC species 3A4 protein. The corresp. gene was amplified from a human liver  
CC derived cDNA library as 2 fragments of 0.6 and 0.9 kb using primers  
CC AAT26933-6. The prod. was cloned into the yeast expression vector pAA45N  
CC to generate plasmid p3A4 for prodn. of the cytochrome only or into the  
CC vector pAHRR to generate the plasmid p3A4R for co-prodn. with the yeast  
CC NADPH-P450 reductase. The sequence is placed under control of the yeast

ADH gene promoter and terminator. The vectors are used in a method for evaluating the safety of a cpd. by reacting the test cpd. with recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380), 2C9 (AAT28381), 2E1 (AAT28382), 3A4 or their variants (AAT28384-98) together with yeast NADPH-P450 reductase (either as a fused protein or as a cell extract) and analysing the resultant metabolite. The cpd. is considered "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it is not detoxified or is metabolised to a carcinogenic cpd

XX Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSW--ASIAFSLYLAPRRS--SLYNLQG--PNHTNY--FTGNFLDILSARTG-----E 61  
 DB 7 LAMETWLLAVSLVLLYLGTHSHGLFKLGIPGTPPLPFLGN--ILSYHKGFCMFDM 63

QY 62 EHAKYREKYGSLRAGIAGAVLNSTDPKVFNVHM--KEAYDY---PKPGMAARVLRIAT 117  
 DB 64 CHKYGKVGWF---YDG--QQPVLAITDPDMIKTVLVKECVSVFTNRRPFGPVGFMKSA- 117

QY 118 GDGVVTAEGEAHKKRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAGES 177  
 DB 118 ---ISIAEDEEWKELRSLSPFTTSGKLKEMVPIIAQYGDVLVRLRREA----- 164

QY 178 AGEKKAATLET-EGVDVKDWGRATLDVMAAGFDYKSDSLQNTKTNLYVAFVGLTDGFA 236  
 DB 165 -----ETGKPVTLKDVFGAYSDMTSTSGVNISSLNPNQD----- 201

QY 237 PTLDSFKAIM-WDFV-PYFRMK-RRHEIPTQGLAV---SRRVGIELMEQKQAVLGS 290  
 DB 202 PFVENTKLLRFDLDPFLSITVPFLIPILEVLNVCVPPREVTNFRKSVKR-----M 256

QY 291 SDQAVDKDVQGRDILSLVLRANIANLPESQKLSDEVLQAIQNLIFAGYETSSVTW 350  
 DB 257 KESRLEDTQKRVDFLQIMIDSQNSKETESHKALSDELVAQSIIFIFAGYETSSVLSF 316

QY 351 MFHRLSEDKAVQDKLREEICQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408  
 DB 317 IMYELATHPDVQKQLEEDAVLPNKAPPTDVLQMEYLDVMVNETLRFPIAMELERV 376

QY 409 CLKDEDFIPLAEFVIGRDGVSINEVRITKGTMMWMLPLFNINRSKFIYGEDAEFRPERWL 468  
 DB 377 CKKQVE-----INGMFIPKGVWVMPISYALHRDP-KYWTPEKFLPERFS 420

QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRRVQFEP 520  
 DB 421 KKNKNDIDPIYITPFG-----SGPRNCIGMRFAVMNKKLALIRVLQNFSGPK 467

RESULT 8  
 AAY05202  
 ID AAY05202 standard; protein; 503 AA.

AC AAY05202;  
 XX  
 XX 17-JUN-1999 (first entry)  
 DT Human CYP3A4 protein.  
 DE CYP3A4 gene polymorphism; polymorphic locus; human; altered metabolism;  
 KW CYP3A4 substrate; drug-drug interaction identification; toxin exposure;  
 KW Genetic linkage detection; phenotypic variation.  
 XX Homo sapiens.  
 XX WO9913106-A1.  
 XX 18-MAR-1999.  
 PD  
 XX 02-SEP-1998; 98WO-US018158.  
 PF

XX 10-SEP-1997; 97US-0058612P.  
 XX (AAYS-) AAYS PHARM INC.  
 PA Lichter JB, Guida M;  
 PI WPI; 1999-215070/18.  
 XX N-PSDB; AAX28295.  
 DR  
 XX New isolated CYP3A4 polymorphic sequences.  
 PT  
 XX Disclosure; Page 25-27; 40pp; English.  
 PS  
 XX This sequence represents the human CYP3A4 gene protein. The invention relates to a CYP3A4 sequence polymorphism, which is part of a non-naturally occurring chromosome. Nucleic acids comprising the CYP3A4 polymorphic sequences can be used to screen patients for altered metabolism for CYP3A4 substrates, potential drug-drug interactions, and adverse/side effects as well as diseases that result from environmental or occupational exposure to toxins. They can also be used to establish animal, cell culture and in vitro cell-free models for drug metabolism. Polymorphic CYP3A4 gene sequences can be used for expression studies to determine the effect of promoter and/or intron sequence variations on mRNA expression and stability. The polymorphisms are also used as single nucleotide polymorphisms to detect genetic linkage to phenotypic variation in activity and expression of CYP3A4. The nucleic acids can also be used to generate genetically modified non-human animals or site specific gene modifications in cell lines

XX Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSW--ASIAFSLYLAPRRS--SLYNLQG--PNHTNY--FTGNFLDILSARTG-----E 61  
 DB 7 LAMETWLLAVSLVLLYLGTHSHGLFKLGIPGTPPLPFLGN--ILSYHKGFCMFDM 63

QY 62 EHAKYREKYGSLRAGIAGAVLNSTDPKVFNVHM--KEAYDY---PKPGMAARVLRIAT 117  
 DB 64 CHKYGKVGWF---YDG--QQPVLAITDPDMIKTVLVKECVSVFTNRRPFGPVGFMKSA- 117

QY 118 GDGVVTAEGEAHKKRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAGES 177  
 DB 118 ---ISIAEDEEWKELRSLSPFTTSGKLKEMVPIIAQYGDVLVRLRREA----- 164

QY 178 AGEKKAATLET-EGVDVKDWGRATLDVMAAGFDYKSDSLQNTKTNLYVAFVGLTDGFA 236  
 DB 165 -----ETGKPVTLKDVFGAYSDMTSTSGVNISSLNPNQD----- 201

QY 237 PTLDSFKAIM-WDFV-PYFRMK-RRHEIPTQGLAV---SRRVGIELMEQKQAVLGS 290  
 DB 202 PFVENTKLLRFDLDPFLSITVPFLIPILEVLNVCVPPREVTNFRKSVKR-----M 256

QY 291 SDQAVDKDVQGRDILSLVLRANIANLPESQKLSDEVLQAIQNLIFAGYETSSVTW 350  
 DB 257 KESRLEDTQKRVDFLQIMIDSQNSKETESHKALSDELVAQSIIFIFAGYETSSVLSF 316

QY 351 MFHRLSEDKAVQDKLREEICQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408  
 DB 317 IMYELATHPDVQKQLEEDAVLPNKAPPTDVLQMEYLDVMVNETLRFPIAMELERV 376

QY 409 CLKDEDFIPLAEFVIGRDGVSINEVRITKGTMMWMLPLFNINRSKFIYGEDAEFRPERWL 468  
 DB 377 CKKQVE-----INGMFIPKGVWVMPISYALHRDP-KYWTPEKFLPERFS 420

QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRRVQFEP 520  
 DB 421 KKNKNDIDPIYITPFG-----SGPRNCIGMRFAVMNKKLALIRVLQNFSGPK 467

RESULT 9	
ABG68753	
ID	ABG68753 standard; protein; 503 AA.
XX	
AC	ABG68753;
XX	
DT	07-OCT-2002 (first entry)
XX	
DE	Cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 2-13.
XX	
KW	Cytochrome P450; CYP3AP1; CYP3AP2; CYP3A43; CYP3A4; CYP3A5; CYP3A7;
KW	drug metabolism; drug design; drug screening.
XX	
OS	Homo sapiens.
XX	
PN	WO200244213-A1.
XX	
PD	06-JUN-2002.
XX	
PF	28-NOV-2001; 2001WO-SE002631.
XX	
PR	28-NOV-2000; 2000SE-00004366.
XX	
PR	11-JUN-2001; 2001SE-00002061.
XX	
PA	(ZAPH/) ZAPHIROPOULOS P G.
PA	(FINT/) FINTA C.
XX	
PI	Zaphiropoulos PG, Finta C;
XX	
DR	WPI; 2002-557532/59.
DR	N-PSDB; ABK97692.
XX	
PT	Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
PT	of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
PT	metabolism, in drug design and drug screening.
XX	
PS	Claim 2; Fig 11; 131pp; English.
XX	
CC	The invention describes a cytochrome P450 protein (I) in which CYP3A43
CC	exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
CC	fragments, variants and multiples of (I) having essentially the same
CC	characteristics. (I) is useful as a medicament, and for evaluating drug
CC	metabolism, in drug design, and drug screening, and in tests for
CC	adjusting the dose of drugs. This is the amino acid sequence of a novel
CC	cytochrome P450 protein
XX	
SQ	Sequence 503 AA;
Query Match	13.8%; Score 394; DB 5; Length 503;
Best Local Similarity	28.7%; Pred. No. 2.1e-25;
Matches 142; Conservative	83; Mismatches 177; Indels 92; Gaps 20;
Qy	47 FTGNFLDILSARTG-----EEHAKYREKYGSTLRPAGIAPVLNSTDPKVFNIHM-KE 99
Db	46 FLGN---ILSYHKGFCWFMDECHKYGVKMGF---YDG--QQFVLAITDPDMIKTVLVE 97
Qy	100 AYDY---PKGMAARVLRIATGQVVTAGEAKHRRIMIPSLSAQAVKSMVPFLFKG 156
Db	98 CYSVFTNRRPFGVGFMKSA-----ISIADEEWKRLRLSLFTTSGKLKEMVPIAQYG 153
Qy	157 MELVDKMDAABKXAVGSSAGEKKATRLT-EGVDVXDWVGRTLDVMAIAGFDYKSD 215
Db	154 DVLVRNLREA-----ETGKPVTLKDVFGAYSMVDVITSTSGVNID 194
Qy	216 SLQKNTNELYVAFGLTDGAPFLDSFKAIM-WDFV-PYFRITMK-RRHEIPLTQGLAV-- 270
Db	195 SLNPNQD-----PVENTKLLRFLDFLDPFLFTVFPFPIPLEVINTICV 240
Qy	271 -SRVGIEMEQKQKQAVLGSSAQVADKDVQGRDILSLVRANIAANLPESQKLSDEEV 329
Db	241 FPREVTFLEKSVKR-----MKESRLDITQKRVDFLQLMIDSNQSKETESHKALSDEL 295
Qy	330 LAQISNLLFAGYETSSVLTWMTFHRUSEKXAVQDKLUREBICQI--DTPMTDLDELNALPY 387
Db	296 VAQSIIFIFAGYETSSVLSFIMYELATHPDVQVKLQEBIDAVLPNKAPPTDVLQMEY 355
Qy	388 LEAFVKESLRIDPPSPYANRECLKDEDFIPLAEPVIGRDSVINEVRIYKGTVMVLPFN 447
Db	356 LDMVYNETLRLFPIAMLERVCKKDVE-----INGMFIPKGVVVMIPSYA 400
Qy	448 INRSKFIYGEDAEERFRERWLEDVTDLSNS-IEAPYGHQASFTSGPRACFGMRFAVAEMK 506
Db	401 LHRDP-KWTEPEXELPERFSKKKNIDPIYITPFG-----SGPRNCIGMRFALNMNK 453
Qy	507 AFLFVTLRVRQFEP 520
Db	454 LALIRVLQNFSEKP 467
RESULT 10	
ABB08079	
ID	ABB08079 standard; protein; 527 AA.
XX	
AC	ABB08079;
XX	
DT	10-SEP-2002 (first entry)
XX	
DE	Maize cytochrome P450, CYP72A1.
XX	
KW	Maize; cytochrome P450; CYP72A1; CYP92A1; plant; gene therapy; antidote.
XX	
OS	Zea mays.
XX	
PH	Key
FT	Peptide
FT	/note= "this fragment is not provided in the protein
FT	sequence indicated in the sequence listing"
XX	
PN	US6380465-B1.
XX	
DD	30-APR-2002.
XX	
PF	11-JUL-1999; 99US-00351229.
XX	
PR	12-JUL-1998; 98US-0092596P.
XX	
PA	(KENT ) UNIV KENTUCKY RES FOUND.
XX	
PI	Barrett M;
XX	
DR	WPI; 2002-470227/50.
DR	N-PSDB; ABL60758.
XX	
PT	Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1,
PT	useful for the production of plants resistant to heterocyclic,
PT	sulfonylurea and substituted urea herbicides and organophosphate
PT	insecticides e.g. Classic and Pursuit.
XX	
PS	Claim 13; Fig 1; 21pp; English.
XX	
CC	The invention relates to isolated nucleic acid molecules encoding Zea
CC	mays (maize) cytochrome P450 designated CYP72A1 and CYP92A1. The
CC	polypeptides can be expressed by standard recombinant methodology. The
CC	nucleic acids may be used in genetic engineering protocols to transform
CC	plants and other eukaryotes e.g. yeast, maize (especially), soybean,
CC	beet, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,
CC	rose, carnation, gerbera, carrot, tomato, lettuce, chicory, pepper, melon
CC	and cabbage. They may be used in this way to confer protection against
CC	heterocyclic herbicides, sulfonylurea herbicides, substituted urea
CC	herbicides and/or organophosphate insecticides. The present sequence
CC	represents the Z. mays CYP72A1
XX	
SQ	Sequence 527 AA;
Query Match	13.8%; Score 392.5; DB 5; Length 527;
Best Local Similarity	24.6%; Pred. No. 3.1e-25;



DT 07-OCT-2002 (first entry)  
XX Cytochrome P450 CYP3A7 and CYP3A1 fusion protein.  
DE  
XX  
KW CYP3A7; CYP3A1; CYP3A2; CYP3A4; CYP3A5; CYP3A7;  
KW cytochrome P450; CYP3A7 and CYP3A1 fusion protein.  
KW drug metabolism; drug design; drug screening.  
XX  
OS Homo sapiens.  
XX  
XX W0200244213-A1.  
XX  
XX PD 06-JUN-2002.  
XX  
XX 28-NOV-2001; 2001WO-SB002631.  
XX  
XX 28-NOV-2000; 2000SE-00004366.  
PR 11-JUN-2001; 2001SE-00002061.  
XX  
XX (ZAPH/) ZAPHIROPOULOS P G.  
PA (FINT/) FINTA C.  
XX  
XX Zaphiropoulos PG, Finta C;  
PI  
XX WPI; 2002-557532/59.  
DR N-ESDB; ABK97643.  
XX  
XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets  
PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug  
PT metabolism, in drug design and drug screening.  
XX  
XX Claim 1; Page 73-75; 131pp; English.  
XX  
XX The invention describes a cytochrome P450 protein (I) in which CYP3A43  
CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub  
CC fragments, variants and multiples of (I) having essentially the same  
CC characteristics (I) is useful as a medicament, and for evaluating drug  
CC metabolism, in drug design, and drug screening, and in tests for  
CC adjusting the dose of drugs. This is the amino acid sequence of a novel  
CC cytochrome P450 protein  
XX  
XX Sequence 535 AA;  
SQ  
Query Match 13.5%; Score 385.5; DB 5; Length 535;  
Best Local Similarity 29.08; Pred. No. 1.3e-24;  
Matches 157; Conservative 79; Mismatches 151; Indels 115; Gaps 26;  
13 LAAPSW--ASTAPFSLYL-APRRSSLYNLOQ-PNHTNY-FTGNFLDLSARTG-----E 61  
7 LAVETWLLAVSLILLYLGYTRHGLPKLGIPOPTPLPLGN---ALGFRKGYWTFDME 63  
62 BHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHVM-KEAYD-----YKFGMAARVL 113  
64 CYKRYKRWG-----IYDCQPMALAITDPMIKTVLKECYSVFTNRRPFGVPVGMKNAI 118  
114 RIATGCVVTAAGBAHRHRIRIMPSLSAQAQVSMVPIFEKGMELVDKMDAAEKDMA 173  
119 SI-----AEDEWKRISLLSPTTSCKLKEMVPIIAQYGDVLVNLREA----- 164  
174 VGESAGEKKAATRLT-EGVDVQWVGWGRATLDVMAAGFDYKSDSLQNKTNELYVAVGLT 232  
165 -----ETGKFVTLKHVFGAYSMDVITSTSGVSDSLNNPOD-----PEVENT 207  
233 D---GFAPTLDSEKAIMWDFVPYERTWKRRIEPLTOGLAVS---RRVGIELMEQKQAV 286  
208 KKLRFNP-LDPFVLSIKVP-PFL-----PILALNITVFPKVIISFLTCKSVKQIK 257  
287 LGSASDAQVDKDVQGR--DILSLVLRANTAAANLPESQKLSDEVLAAQISNLLFAGYETS 344  
258 EGRL-----KETQKRVDFLQMLDMSQNSKDSETHKALSDELMAQSIIFIFAGYETT 310  
345 STVLTWTHRLSDEKAVQDKLREICQIDTDM-----PTLDENALPYLEAFYKESLRLD 399  
311 SSVLSFIITIELATHPDVQKQVKE---IDTVLPNKAPPTVDTVQLQLEYLDMVNVNLRLP 367

QY 400 PPSVYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMMLPLFNINRSKFIYGEDA 459  
DB 368 PYAMRLERVCKDVE-----INGMFIKGVVVVMPISVYLHDP-KYWTEP 411  
QY 460 EBFRRPERMLEVDTSLS-IBAPYHQASFIISGRACFGWFAVAEMKAFILVTLRRVQF 518  
DB 412 EKFLPERFSKKNKNDIPYITPFG-----SGPRNCIGMRFALVNMKALVRLVQNF 465  
QY 519 EP 520  
DB 466 KP 467

RESULT 13  
AAE26192  
ID AAE26192 standard; protein; 502 AA.  
XX  
XX AAE26192;  
XX  
XX 14-NOV-2002 (first entry)  
XX Human cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) protein.  
XX  
XX Human; cytochrome P450; subfamily IIIA; polypeptide 5 isogene; CYP3A5;  
KW drug screening; polymorphism; haplotype; drug metabolising disorder;  
KW gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 30 /note= "This amino acid changes to Tyr due to single  
XX nucleotide polymorphism"  
XX Misc-difference 100 /note= "This amino acid changes to Tyr due to single  
XX nucleotide polymorphism"  
XX  
XX W0200246209-A2.  
XX  
XX 13-JUN-2002.  
XX  
XX 07-DEC-2001; 2001WO-US047218.  
XX  
XX 08-DEC-2000; 2000US-0254367P.  
PR 03-MAY-2001; 2001US-0288470P.  
XX  
XX (GENA-) GENAISSANCE PHARM INC.  
XX  
XX Anastasio AE, Han J, Klien SE, Rounds E;  
XX  
XX WPI; 2002-636448/68.  
XX N-PSDB; AAD43350, AAD43351.  
XX  
XX Novel isolated polynucleotide which is a polymorphic variant of  
XX cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) gene useful for  
XX expressing CYP3A5 protein isoform used in drug screening techniques.  
XX  
XX Claim 29; Fig 3; 127pp; English.  
XX  
XX The invention relates to isolated polynucleotide having cytochrome P450,  
XX subfamily IIIA, polypeptide 5 isogene (CYP3A5). The invention is useful  
XX for screening drugs. The invention is useful for studying expression and  
XX function of CYP3A5 and expressing CYP3A5 protein for use in screening for  
XX candidate drugs to treat diseases related to CYP3A5 activity. The  
XX polymorphism and haplotype data is useful for validating whether CYP3A5  
XX is a suitable target for drugs to treat drug metabolising disorders.  
XX screening for such drugs and reducing bias in clinical trials of such  
XX drugs. The invention is also useful for therapeutic purposes. The  
XX invention is useful in studying the effect of variation on the biological  
XX activity of CYP3A5 as well as on the binding affinity of candidate drugs  
XX to CYP3A5, or for studying the enzymatic properties of such CYP3A5  
XX variants using these candidate drugs as substrate. The invention is



CC useful in gene therapy. The present sequence is human CYP3A5 protein  
 XX  
 SQ Sequence 502 AA;  
 Query Match 13.5%; Score 385; DB 5; Length 502;  
 Best Local Similarity 28.1%; Pred. No. 1.3e-24;  
 Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;  
 QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOQ-PNHTNY-FTGNFLDILSARTG-----E 61  
 Db 7 LAVETWLLAVSLVLLYLYGTRTHGLFKRLGIPGTPPLPLGN---VLSYRQGLWKEDTE 63  
 QY 62 EHAKYREKYGSTRFAGIAGAPVLNSTDPKVFNHVM-KEAYDYPKPGMAARVLRATG-- 118  
 Db 64 CYKYGKMGWGT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111  
 QY 119 ----DGVVTAEGEAHKHRRIMIPSLSAQAVKSNVPIFLEKGMELVDKMDAEAKDMAV 174  
 Db 112 GFMKSAISLAEDDEWKRIIRSLSPFTTSGLKEMFPPIIAQYGDVLRVLRRE-AEKG--- 167  
 QY 175 GESAGEKKATRLTEGVVDKVMGVRATLDVVALAGFDYKSDSLQN-----KTNELVYAF 228  
 Db 168 -----KPVTLKDIFGAYSMVDVITGTSFGVNIIDSLNNPQDPFVSTKFLKF 213  
 QY 229 VGLTGFAPTLDSFKAIMWDFVYFRMKRHEIPLTQGLAVS--RRVGLMEQKKQAV 286  
 Db 214 -----GFLDPL-FLSIILFPFL-----TFVFEALNVSLFPKDTINFLSKS--- 252  
 QY 287 LGSASDAQVDKQVQGRDILSLVRANIANLPESQKLSDEEVLAAQISNLLFAGYETSST 346  
 Db 253 VNRMKSRNLNDKQKRLDFLQIMIDSQNSKETESHKALSDELAQAQSIIFIFAGYETSS 312  
 QY 347 VLTWMPHRLSDEKAVQDKLREIQQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPY 404  
 Db 313 VLSFTLYELATHDPVQOKLQKEIDAVLPNKAPPTYDAVVQMEYLDVMVNETLRLFPVAIR 372  
 QY 405 ANRECLKDEDFFLAEPVIGRDSVINEVRIITKGTMMWMLPLFNINRSFIYGEDAEFRP 464  
 Db 373 LERTCKRDVE-----INGVFIKGSWVPIYALHNDP-KYWTPEFERP 416  
 QY 465 ERWLEDVTSLSN-IEAPYGHQASFTSGPRACFGMRFAVEMKAPFLVTLRRVQFEP 520  
 Db 417 ERSK-KDSDIDPIYITPFG-----TGPRNCIGMRFALNMKALIRVLQNFSPK 466  
 RESULT 14  
 ADE57186  
 ID ADE57186 standard; protein; 502 AA.  
 XX AC ADE57186;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human Protein P20815, SEQ ID NO 3047.  
 XX KW Human; pain; neuronal tissue; gene therapy;  
 XX KW spinal segmental nerve injury; chronic constriction injury;CCI;  
 XX KW spared nerve injury; SNI; Chung.  
 XX OS Homo sapiens.  
 XX FN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX FF 14-AUG-2002; 2002WO-US025765.  
 XX FR 14-AUG-2001; 2001US-0312147P.  
 XX PR 01-NOV-2001; 2001US-0346382P.  
 XX PR 26-NOV-2001; 2001US-0333347P.  
 XX PR (GCHO ) GEN HOSPITAL CORP.  
 XX PA (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;  
 PI WPI; 2003-268312/26.  
 XX GENBANK; P20815.  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 502 AA;  
 SQ  
 Query Match 13.5%; Score 385; DB 7; Length 502;  
 Best Local Similarity 28.1%; Pred. No. 1.3e-24;  
 Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;  
 QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOQ-PNHTNY-FTGNFLDILSARTG-----E 61  
 Db 7 LAVETWLLAVSLVLLYLYGTRTHGLFKRLGIPGTPPLPLGN---VLSYRQGLWKEDTE 63  
 QY 62 EHAKYREKYGSTRFAGIAGAPVLNSTDPKVFNHVM-KEAYDYPKPGMAARVLRATG-- 118  
 Db 64 CYKYGKMGWGT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111  
 QY 119 ----DGVVTAEGEAHKHRRIMIPSLSAQAVKSNVPIFLEKGMELVDKMDAEAKDMAV 174  
 Db 112 GFMKSAISLAEDDEWKRIIRSLSPFTTSGLKEMFPPIIAQYGDVLRVLRRE-AEKG--- 167  
 QY 175 GESAGEKKATRLTEGVVDKVMGVRATLDVVALAGFDYKSDSLQN-----KTNELVYAF 228  
 Db 168 -----KPVTLKDIFGAYSMVDVITGTSFGVNIIDSLNNPQDPFVSTKFLKF 213  
 QY 229 VGLTGFAPTLDSFKAIMWDFVYFRMKRHEIPLTQGLAVS--RRVGLMEQKKQAV 286  
 Db 214 -----GFLDPL-FLSIILFPFL-----TFVFEALNVSLFPKDTINFLSKS--- 252  
 QY 287 LGSASDAQVDKQVQGRDILSLVRANIANLPESQKLSDEEVLAAQISNLLFAGYETSST 346  
 Db 253 VNRMKSRNLNDKQKRLDFLQIMIDSQNSKETESHKALSDELAQAQSIIFIFAGYETSS 312  
 QY 347 VLTWMPHRLSDEKAVQDKLREIQQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPY 404  
 Db 313 VLSFTLYELATHDPVQOKLQKEIDAVLPNKAPPTYDAVVQMEYLDVMVNETLRLFPVAIR 372

QY 405 ANRECKDEDFIPLAEFVIGRDSVINEVIRKGTWMLPLFNINRSKFIYGEDAEERFP 464  
Db 373 LERTCKRDE-----INGVFIKGSVMVPTIALHHDH-KYWTPEERFP 416  
QY 465 ERWLEDVTDLSNS-IBAPYGHQASFIISGPRACFGWRFVAEMKAFLEFVTLRRVQEP 520  
Db 417 ERSKX-KDSIDPVIYTPFG-----TGPRNCIGMRFAIMNMKALIRVLQNFSEKP 466

RESULT 15  
ADE57190  
ID ADE57190 standard; protein; 502 AA.  
XX  
AC ADE57190;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P20815, SEQ ID NO 3051.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
WO2003016475-A2.  
XX  
27-FEB-2003.  
XX  
14-AUG-2002; 2002WO-US025765.  
XX  
14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
PI  
PI Woolf C, D'urso D, Befort K, Costigan M;  
DR WPI; 2003-268312/26.  
DR GENBANK; P20815.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 502 AA;  
Query Match 13.5%; Score 385; DB 7; Length 502;  
Best Local Similarity 28.1%; Pred. No. 1.3e-24;  
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;  
QY 13 LAAPSW--ASIAFFSLYL-APRSSLYNLOG-PNHNTY-FTGNFLDILSARTG-----E 61  
Db 7 LAVETWLLAVSLVLLVLYGTHGLFRKLGIPGPTPLPLGN---VLSYQGLWKEDTE 63  
QY 62 ERAKYREKYGSTIRFAGIAGAPVLSNTPKVFNVHM-KEAYDYPKPGMAARVLIATG-- 118  
Db 64 CYKKGKMWGT--YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRSRLGPV 111  
QY 119 -----DGVTTAEAEAKHRRIMIPSLSAQAVKSMVPIFEKGMELVDKMDAEKDMAV 174  
Db 112 GFVKSAISLAEDDEWKIRNSLLSPFTSGKLKEMFPIIAQYGDVLVRLRE-AEKG--- 167  
QY 175 GESAGEKKATRETEGVDVKDWVGRATLDVMALAGFDYKSDSLQN-----KTNELYVAF 228  
Db 168 -----KPVTLKDIFGAYSMVDVITGTSFGVNIDSLNNPQDPFVESTKKPLKF 213  
QY 229 VGLTDGFAPTLDSFKAIMWDFVYPFTWKRRHEIPLTQGLAVS--RRVGIELMEQKKQAV 286  
Db 214 -----GFLDPL-FLSILFPFL-----TFVFEALNVSLFPKDTINFLSKS---- 252  
QY 287 LGSASDAQVKKDVQGRDILSLVRAANTAAANLPEQSKLSDEEVLQAISNLLFAGYETSST 346  
Db 253 VNRKYSRLNDKQKRLDPLQLMIDSONSKETESHKALSDLELAQSIIFIFAGYEITSS 312  
QY 347 VLTWFMHRLSEDKAVQDKLREICQI--DTDMPTLDELNALPYLEAFVKESIRLDPPSPY 404  
Db 313 VLSFTLYELATHPDVQCKLQKEIDAVLPNKAPPTYDAVVQMEYLDVMVNETLRLFPVAIR 372  
QY 405 ANRECKDEDFIPLAEFVIGRDSVINEVIRKGTWMLPLFNINRSKFIYGEDAEERFP 464  
Db 373 LERTCKRDE-----INGVFIKGSVMVPTIALHHDH-KYWTPEERFP 416  
QY 465 ERWLEDVTDLSNS-IBAPYGHQASFIISGPRACFGWRFVAEMKAFLEFVTLRRVQEP 520  
Db 417 ERSKX-KDSIDPVIYTPFG-----TGPRNCIGMRFAIMNMKALIRVLQNFSEKP 466

Search completed: April 2, 2004, 13:58:35  
Job time : 58 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 10:03:02 ; Search time 717 Seconds

(without alignments)  
10031.675 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932  
Sequence: 1 gaattcgccagcagccagccaccc.....agccgcgcgcgcgcgaattc 1932

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCRUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1932	100.0	1932	14	US-10-066-007-2
2	353.2	18.3	3969	14	US-10-066-007-4
3	49.8	2.6	467	12	US-10-424-599-20949
4	49.4	2.6	1501	15	US-10-264-237-392
5	49.2	2.5	368	12	US-10-424-599-1793
6	48.4	2.5	657	12	US-10-424-599-25496
7	48.4	2.5	1512	15	US-10-313-863A-55
8	48.4	2.5	2011	9	US-09-880-107-1586
9	48.4	2.5	2759	14	US-10-146-575-1
10	48.4	2.5	2768	14	US-10-268-822-14
11	48.2	2.5	2849	9	US-09-880-107-2110
12	48.2	2.5	312	9	US-09-960-357-4611
13	47.6	2.5	3059	9	US-09-925-297-268
14	47.4	2.5	628	12	US-10-424-599-71313
15	47.2	2.4	1228	12	US-10-424-599-32438

C 16	47	2.4	218	9	US-09-933-797-162	Sequence 162, App
C 17	47	2.4	685	12	US-10-424-599-8656	Sequence 8656, App
C 18	46.8	2.4	709	12	US-10-424-599-75080	Sequence 75080, A
C 19	46.8	2.4	1151	12	US-10-424-599-23175	Sequence 23175, A
C 20	46.8	2.4	1712	9	US-09-981-876-106	Sequence 106, App
C 21	46.8	2.4	1712	10	US-09-148-876-106	Sequence 106, App
C 22	46.8	2.4	1822	9	US-09-981-876-105	Sequence 105, App
C 23	46.8	2.4	1822	10	US-09-148-876-105	Sequence 105, App
C 24	46.6	2.4	1091	12	US-10-424-599-33549	Sequence 33549, A
C 25	46.4	2.4	1419	12	US-10-424-599-72966	Sequence 72966, A
C 26	46.2	2.4	672	12	US-10-424-599-99744	Sequence 106636, A
C 27	46.2	2.4	710	12	US-10-424-599-99744	Sequence 99744, A
C 28	46.2	2.4	1194	12	US-10-424-599-15131	Sequence 15131, A
C 29	46	2.4	266	12	US-10-424-599-29930	Sequence 29930, A
C 30	46	2.4	414	12	US-10-424-599-44827	Sequence 44827, A
C 31	46	2.4	580	12	US-10-424-599-94220	Sequence 94220, A
C 32	45.8	2.4	592	14	US-10-066-543-850	Sequence 111262, A
C 33	45.8	2.4	991	12	US-10-424-599-54757	Sequence 54757, App
C 34	45.8	2.4	1498	12	US-10-424-599-54758	Sequence 54758, A
C 35	45.8	2.4	2301	12	US-10-424-599-138418	Sequence 138418, A
C 36	45.6	2.4	447	12	US-10-424-599-142657	Sequence 142657, A
C 37	45.6	2.4	583	12	US-10-424-599-52047	Sequence 52047, A
C 38	45.6	2.4	705	12	US-10-424-599-137273	Sequence 137273, A
C 39	45.6	2.4	733	12	US-10-424-599-118266	Sequence 118266, A
C 40	45.6	2.4	558	12	US-10-424-599-67477	Sequence 67477, A
C 41	45.4	2.3	662	12	US-10-424-599-80087	Sequence 80087, A
C 42	45.4	2.3	767	12	US-10-424-599-43674	Sequence 43674, A
C 43	45.4	2.3	1353	15	US-10-369-493-141849	Sequence 141849, A
C 44	45.4	2.3	2508	12	US-10-424-599-141849	Sequence 141849, A
C 45	45.4	2.3				

## ALIGNMENTS

RESULT 1  
US-10-066-007-2  
Sequence 2, Application US/10066007

Publication No. US2003077691A1

GENERAL INFORMATION: Tatsuo

APPLICANT: HOSHINO, Tatsuo

APPLICANT: OJIMA, Kazuyuki

APPLICANT: SEROGUCHI, Yutaka

TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE

FILE REFERENCE: ASTAXANTHIN SYNTHETASE

CURRENT APPLICATION NUMBER: US/10/066,007

CURRENT FILING DATE: 2001-02-01

PRIOR APPLICATION NUMBER: US/09/518,386

PRIOR FILING DATE: 2000-03-03

PRIOR APPLICATION NUMBER: EP 99104668.1

PRIOR FILING DATE: 1999-03-09

PRIOR APPLICATION NUMBER: EP 00101666.6

PRIOR FILING DATE: 2000-02-01

NUMBER OF SEQ ID NOS: 32

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 1932

TYPE: DNA

ORGANISM: Phaffia rhodozyma

FEATURE:

NAME/KEY: CDS

LOCATION: (33)..(1706)

NAME/KEY: POLYA\_site

LOCATION: (1871)

NAME/KEY: mRNA

LOCATION: (14)..(1891)

US-10-066-007-2

Query Match 100.0%; Score 1932; DB 14; Length 1932;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GAATTCGCGCAGGCGCAGCCACTTCTCCATATGTCATCTTGTGCTGTCACAGTGG 60



PRIOR APPLICATION NUMBER: EP 99104668.1  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: EP 00101666.6  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 4  
LENGTH: 3969  
TYPE: DNA  
ORGANISM: Phaffia rhodozyma  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (517)..(518)  
NAME/KEY: intron  
LOCATION: (784)..(898)  
NAME/KEY: intron  
LOCATION: (1016)..(1087)  
NAME/KEY: intron  
LOCATION: (1180)..(1302)  
NAME/KEY: intron  
LOCATION: (1518)..(1600)  
NAME/KEY: intron  
LOCATION: (1635)..(1723)  
NAME/KEY: intron  
LOCATION: (1867)..(1939)  
NAME/KEY: intron  
LOCATION: (2000)..(2081)  
NAME/KEY: intron  
LOCATION: (2182)..(2257)  
NAME/KEY: intron  
LOCATION: (2355)..(2431)  
NAME/KEY: intron  
LOCATION: (2543)..(2618)  
NAME/KEY: intron  
LOCATION: (2653)..(2742)  
NAME/KEY: intron  
LOCATION: (2815)..(2962)  
NAME/KEY: intron  
LOCATION: (3051)..(3113)  
NAME/KEY: intron  
LOCATION: (3172)..(3247)  
NAME/KEY: intron  
LOCATION: (3322)..(3398)  
NAME/KEY: intron  
LOCATION: (3424)..(3513)  
NAME/KEY: polyA site  
LOCATION: (3865)..(3866)  
NAME/KEY: intron  
LOCATION: (653)..(734)  
US-10-066-007-4

Query Match 18.3%; Score 353.2; DB 14; Length 3969;  
Best Local Similarity 99.2%; Pred. No. 1.1e-98;  
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1514 TTGCTTTGGTGGCGATTCGTCGCGAGATGAGCGCTTCTTGTCTCACTCCCG 1573  
DB 3508 TCGCAGTGGTGGCGATTCGTCGCGAGATGAGCGCTTCTTGTCTCACTCCCG 3567  
QY 1574 TCGGGTCCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACATCACCTTGATCAT 1633  
DB 3568 TCGGGTCCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACATCACCTTGATCAT 3627  
QY 1634 TTCCCGTCTCGAATCGTTGGTAGAGAGAGAGGGGTACCAGATGGTTGCGAGTCAA 1693  
DB 3628 TTCCCGTCTCGAATCGTTGGTAGAGAGAGAGGGGTACCAGATGGTTGCGAGTCAA 3687  
QY 1694 GCGGTCGGAATGAGTTGATTCCTATATGTTAAGAGAGAGTCTATATCTGAGAGATGTGTG 1753  
DB 3688 GCGGTCGGAATGAGTTGATTCCTATATGTTAAGAGAGAGTCTATATCTGAGAGATGTGTG 3747  
QY 1754 ACTAGACAATGCTTTTGTATGATTTGTTTCTCATCCCGGGGAGGGGCTATGACT 1813

Db 3748 ACTAGACAATGCTTTTGTATGATTTGTTTCTCATACCCGGCAGGCGCTATGACT 3807  
QY 1814 TTACGCTCTATCGTCTGCTCTGACTCTCTTCTTACCTATATATTTATTCATCCG 1871  
DB 3808 TCTACGCTCTATCGTCTGCTCTGACTCTCTTACCTATATATTTATTCATCCG 3865  
RESULT 3  
US-10-424-599-20949  
; Sequence 20949, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 20949  
; LENGTH: 467  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_118922C.1  
US-10-424-599-20949

Query Match 2.6%; Score 49.8; DB 12; Length 467;  
Best Local Similarity 96.2%; Pred. No. 0.00018;  
Matches 51; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1861 TATTCCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGA 1913  
DB 401 TATTCAATCCCAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGA 453

RESULT 4  
US-10-264-237-392  
; Sequence 392, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: Patent in Ver. 3.1  
; SEQ ID NO 392  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: n equals a.t.g, or c  
US-10-264-237-392

Query Match 2.6%; Score 49.4; DB 15; Length 1501;  
Best Local Similarity 63.2%; Pred. No. 0.0005;  
Matches 74; Conservative 1; Mismatches 42; Indels 0; Gaps 0;  
QY 1785 TTTCCTACATCCCGGCGAGGCTATGACTTCTTACGTCTGTCGCTCTGACTCTC 1844  
DB 1370 TTTCCTGCTTGCATCTCTTGTGTACAGACTGCATAGTGTGACGCTATC 1429





	Query Match	2.5%	Score 48.2	DB 9	Length 312
	Beat Local Similarity	55.8%	Pred. No. 0.00044		
	Matches 92	Conservative	0	Mismatches 73	Indels 0
				Gaps 0	
QY	1745	GAATGTGTGACTAGGACAATGCCCTTCTTTGTATCGATTTGTTTCTATACCCGGGAGCC	1804		
DB	175	GAATCAGGAGGAAACTGAGACCTCTAGAGTATTAAATGTAGTTCAAAACCTAGCCATA	116		
QY	1805	GCATGACTTCTACCTCGCTCTATCGCTCGCTCTCTCTTACCCCTATATATTATT	1864		
DB	115	TGATAGACTGATATCATGTATACCTGGACTTTTCTGCTTTAAATAAAAAACCATACGATT	56		

Query Match	2.5%	Score 47.4	DB 12	Length 628
Best Local Similarity	89.5%	Pred. No. 0.0012	6	Indels 0
Matches 51	Conservative	0	Mismatches	Gaps 0
QY	1858	TATTATTCATCCGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGTCCGAG	1914	
ph	61	TATTATCATCGCGCGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGTCTTAG	5	

## RESULT 15



```
US-10-424-599-32438/c
; Sequence 32438, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 32438
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129297C.1
US-10-424-599-32438

Query Match      2.4%; Score 47.2; DB 12; Length 1228;
Best Local Similarity 86.7%; Pred. No. 0.0021; 8; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 8;

Oy 1855 ATATATTATTCCATCCGAAAAAAGCGCGCGCTCGAG 1914
    |||||
Db 68 ATGTGCTTCCCGCCCAAAAAAAGCGCGCGCTCTAG 9

Search completed: March 31, 2004, 12:47:06
Job time : 723 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 54 Seconds  
(without alignments)  
2914.425 Million cell updates/sec

Title: US-10-066-007-3  
Perfect score: 2852  
Sequence: 1 MFLVLTGALGLAAPSWS.....RIVGREKGYQNRLOQKPVPE 557

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	3	AA08713 Amino aci
2	398.5	14.0	503	6	ABU57260 Human chy
3	398.5	14.0	503	6	ABR82024 Human cyt
4	396.5	13.9	502	7	ADD48378 Human Pro
5	396.5	13.9	503	2	AA72363 Human cyt
6	396.5	13.9	503	2	AA81464 Human der
7	396.5	13.9	503	2	AA93170 Human cyt
8	396.5	13.9	503	2	AA05202 Human CYP
9	394	13.8	503	5	ABG68753 Cytochrom
10	392.5	13.8	527	5	ABH08079 Maize cyt
11	385.5	13.5	503	5	AAE22852 Human cyt
12	385.5	13.5	535	5	ABG68747 Cytochrom
13	395	13.5	502	5	AAE26192 Human cyt
14	385	13.5	502	7	AAE57186 Human Pro
15	385	13.5	502	7	AAE57190 Human Pro
16	385	13.5	502	7	AAE63919 Human Pro
17	382	13.4	454	5	ABG68754 Cytochrom
18	373.5	13.1	454	6	ADA48312 Rice prot
19	370	13.0	504	7	AAE57188 Rat Prote
20	370	13.0	504	7	AAE63917 Rat Prote
21	370	13.0	504	7	AAE57184 Rat Prote
22	367	12.9	520	3	ABE23917 Arabidops
23	365	12.8	520	3	ABU54571 Human NOV
24	363.5	12.7	512	5	AAU97096 Abscicisic
25	362.5	12.7	465	3	AA954205 Arabidops

## ALIGNMENTS

### RESULT 1

AA08713  
ID AAB08713 standard; protein; 557 AA.

XX AAB08713;

AC AAB08713;

XX AAB08713;

DT 12-SEP-2003 (revised)

DT 02-JAN-2001 (first entry)

XX Amino acid sequence of an astaxanthin synthetase polypeptide.

XX Astaxanthin synthetase; astaxanthin; beta-carotene; carotenogenic yeast;

KW antioxidant; cancer; colouring reagent; farmed fish; salmon.

XX Xanthophyllomyces dendrorhous.

XX EP1035206-A1.

XX 13-SEP-2000.

XX 03-MAR-2000; 2000EP-00104430.

XX 09-MAR-1999; 99EP-00104668.

XX 01-FEB-2000; 2000EP-00101666.

XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.

XX Hoshino T, Ojima K, Setoguchi Y;

XX WPI; 2000-559874/52.

XX N-PSDB; AAA64472, AAA64473.

XX Novel polynucleotide encoding astaxanthin synthase useful for producing

XX recombinant cells for producing astaxanthin from beta-carotene.

XX Claim 2; Page 21-23; 46pp; English.

XX The present sequence represents an astaxanthin synthetase polypeptide of

XX Phaffia rhodozyma. The enzyme is involved in the last step of the

XX astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.

XX rhodozyma is a carotenogenic yeast strain. The astaxanthin synthetase

XX polynucleotides and polypeptides are useful for producing astaxanthin.

XX Astaxanthin is an antioxidant which may be used to protect living cells

XX against diseases such as cancer. Astaxanthin is also used as a colouring

XX reagent, e.g. in farmed fish like salmon to impart an orange-red

XX coloration. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 557 AA;

XX SQ



RESULT 3  
ID ABR82024  
XX ABR82024 standard; protein; 503 AA.  
AC ABR82024;  
XX 22-SEP-2003 (first entry)  
DT  
XX Human cytochrome P450 3A4 amino acid sequence.  
DE  
XX Human; protein array; protein moiety; phenotype; drug discovery;  
KW naturally occurring variant; pharmacogenomic; diagnostic;  
KW parallel analysis; tumour suppressor; p53; cytochrome P450.  
XX  
XX Homo sapiens.  
XX WO2003048768-A2.  
XX 12-JUN-2003.  
XX 05-DEC-2002; 2002WO-GB005499.  
XX 05-DEC-2001; 2001US-0335806P.  
XX 16-SEP-2002; 2002US-0410815P.  
XX (SENS-) SENSE PROTEOMIC LTD.  
PA  
XX Boutell JM, Godber BLJ, Hart DJ, Blackburn JD;  
XX WPI; 2003-569063/53.  
XX N-PSDB; ACF06054.  
XX  
XX New protein array, useful for determining the phenotype of a naturally  
PT occurring variant of a DNA sequence of interest, comprises a surface upon  
PT which at least two protein moieties are deposited.  
XX  
XX Example 5; Fig 11B; 84pp; English.  
XX  
XX The present invention describes a protein array comprising a surface upon  
CC which at least two protein moieties are deposited at spatially defined  
CC locations, where the protein moieties are naturally occurring variants of  
CC a DNA sequence of interest. Also described: (1) making a protein array;  
CC (2) screening a set of protein moieties for molecules that interact with  
CC one or more proteins; and (3) simultaneously determining the relative  
CC properties of members of a set of protein moieties. The protein array can  
CC be used for determining the phenotype of a naturally occurring variant of  
CC a DNA sequence of interest. The protein array is useful for drug  
CC discovery, pharmacogenomics and diagnostics. The protein array allows the  
CC parallel analysis of closely related proteins with a sensitivity that is  
CC at least comparable to existing methods, if not better, with small  
CC volumes of potentially expensive ligands, and in a quantitative,  
CC comparative, functional analysis manner not previously possible. ACF06000  
CC to ACF06056 and ABR81975 to ABR82025 represent sequences used in the  
XX exemplification of the present invention  
XX  
SQ Sequence 503 AA;  
Query Match 14.0%; Score 398.5; DB 6; Length 503;  
Best Local Similarity 28.7%; Pred. No. 8.6e-26;  
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
13 LAATSW--ASTAFSLYLAPRS--SLYNLOG--PNHTNY--FTGNFLILSARGT-----E 61  
DB 7 LAMETWLLAVSLVLLLYLTHGSHGLFKLGIQPTPLPLGN---ILSYHKGFCNFDME 63  
62 EHAKYREKYGSTLRFAGIAGAPLVNSTDPKVFNHVM--KEAYDY---PKPQMAARVLRIAT 117  
DB 64 CHKKYKVGWGF---YDG--QQVLAIIDPDMIKTVLVKCYSVFTNRPPGPGFMKSA- 117  
118 GDGVVTAPEAKHRRRIIMPSLSAQAVKGNVPIFEKGMELVDKMMEDAAEKDVAVGES 177  
DB 118 ---ISIAEDDEWKRLRLSLSPFTTSGLKEMVPIAQYGDVLVRLNRREA----- 164

QY 178 AGEKKAETLET--EGVDVKDWVGRATLDVMAAGFDYKSDSLQNKTNELYVAFLVGLTDGFA 236  
DB 165 -----ETGKPVTLKDVFGAYSMVDITSTSGVAINDSLNNPD----- 201  
QY 237 PILDSEKAIM-WDFV-PYERTMK-RRHEIPLTQGLAV---SRRVGIELMEQKQOAVLGS 290  
DB 202 PFVENTKLLRDFDLDFPFLSITVFPFLIPILEVNICVFPREVTNFKSVKR-----M 256  
QY 291 SDQAVDKDQVGRDILSLVRANTAAANLPESQKLSDEVLQAQINLLFAGYETSTVLTW 350  
DB 257 KESRLEDTKQHRVDFLQMLMDSQNSKETESHKALSLELVAQSIIFIPAGYETTSVLSF 316  
QY 351 MHRRLSEKAVQDKLREICQI--DTDMPTLDLNAFLYEAFAVKESLRLDPPSPYANRE 408  
DB 317 IMYELATHPDVQOQKLEIDAVLPNKPAPPTTDTVLQMEYLDVWVNEILRLFPIMRLERV 376  
QY 409 CLKDEDFIPLAEFVIGRDSVINEVRIITKGTVMMLPFLNINRSKFIYGEDAEERFRPRL 468  
DB 377 CKXQVE-----INGMFIPKGVVVMIPSYALHRDP-KYWTEPEKELPERFS 420  
QY 469 EVDVTSLSN-IEAPYGHQASFISSGRACFGWRFAVAEMKAPLFVTLRRVQFEP 520  
DB 421 KKKNDNIDPIYIYTPFG-----SGPRNCIGMRFAFMNKKLALIRVLQNFSEKP 467  
RESULT 4  
ADD48378  
ID ADD48378 standard; protein; 502 AA.  
XX  
XX ADD48378;  
DT 29-JAN-2004 (first entry)  
XX Human Protein A29815, SEQ ID NO 14079.  
DE  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
XX Homo sapiens.  
XX W02003016475-A2.  
XX 27-FEB-2003.  
XX 14-AUG-2002; 2002WO-US025765.  
XX 14-AUG-2001; 2001US-0312147P.  
XX 01-NOV-2001; 2001US-0346382P.  
XX 26-NOV-2001; 2001US-0333347P.  
XX (GEHO ) GEN HOSPITAL CORP.  
XX (FARB ) BAYER AG.  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX GENBANK; A29815.  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 101pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 502 AA;

Query Match 13.9%; Score 396.5; DB 7; Length 502;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSW--ASIAFFSLLAPRRS--SLYNLQG--PNHTNY--FTGNFIDLISARTG-----E 61  
 DB 6 LAMETWLLAVSLVLLYLGTHSHGLFKLGPPTPLPFLGN---ILSYHKGFCFMDME 62

QY 62 EHAKEKYGSTLRPAGIAGAPVLNSTDPKFNHYM--KEAYDY---PKPGMAARVLRIAT 117  
 DB 63 CHKKYKVGWGF---YDG--QQPVLAITDPMIKTVLVKECYSVFTNRPPFGVGFMSKA- 116

QY 118 GDGVVTAEGEAHKKRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDMAYGES 177  
 DB 117 ---ISIAEDEWKRLRSLLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 163

QY 178 AGEKKAATLET--EGVDVKDWVGRATLDVMAAGFDYKSDSLQNTKNELYVAFVGLTDGFA 236  
 DB 164 -----ETGKPVTLKDVFGAYSMDVITSTSGVNIIDSLNPNQD----- 200

QY 237 PTLDSFKAIM-WDFV--PYFRTWK--RRHEIPLTQGLAV---SRRVGIELMEQKQAVLGSA 290  
 DB 201 PFVENTKKLLRDFDLPFPFLSITVPFPIILEVLNLCVFPREVTNFRKSVKR-----M 255

QY 291 SDQAVDKKDVQGRDILSLVRANIANLPESQKLSDEEVLQISNLLPAGYETSTVLTW 350  
 DB 256 KESRLEDTKHRVDFLQMLDMSQNSKETESHKALSDELVAQSIIIFAGYETTSVLSF 315

QY 351 MFLRLSDRAVDKLRREEICQI--DTPMTLDLALPYLEAFVKESLRLDPPGPPYANRE 408  
 DB 316 IMYELATHDPVQKQLEEDAVLPKAPPTVTLQMEYLDVYVNETLRLPFIARLERV 375

QY 409 CLKDEDFIPLAEPVIGRDSVINEVRIKGTVMVMLPLFNINRSKFTYGEDAEERPRRWL 468  
 DB 376 CKKDVE-----INGMEIPKGVVMVPIPSALHRDP--KYWTEPEKTLPRFS 419

QY 469 EDVTDLSINS--IEAPYGHQASFTSGPRACFGWFAVAEMKARLFTVLRVQFEP 520  
 DB 420 KXNKNDIPVIYTPFG-----SGPRNCIGMRPALNWKMLALIRVLQNFSPK 466

RESULT 5  
 AAR72363  
 ID AAR72363 standard; protein; 503 AA.  
 XX  
 AC AAR72363;  
 XX  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 14-NOV-1995 (first entry)  
 XX  
 XX Human cytochrome P450 molecular species 3A4 protein.  
 DE  
 XX

KW Human cytochrome P450; amplification; PCR; primer; expression vector;  
 KW yeast NADPH-P450 reductase, safety; fusion protein; metabolite;  
 KW carcinogen; mutagen; liver metabolism.  
 XX Homo sapiens.  
 PN EP644267-A2.  
 XX  
 PD 22-MAR-1995.  
 XX  
 XX 20-JUL-1994; 94EP-00111298.  
 XX  
 XX 20-JUL-1993; 93JP-00201120.  
 PR 21-JUL-1993; 93JP-00180246.  
 PR 30-JUL-1993; 93JP-00208279.  
 XX  
 XX (SUMO ) SUMITOMO CHEM CO LTD.  
 PA  
 XX  
 PI Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;  
 DR WPI; 1995-116991/16.  
 XX  
 XX N-PSDB; AAQ87717.  
 DR  
 XX  
 PT Evaluation of safety of a chemical cpd. - using recombinant Yeast  
 PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.  
 XX  
 XX Example; Page 31-33; 124pp; English.  
 PS  
 XX  
 CC The amino acid sequence of the human cytochrome P450 species 3A4. The  
 CC cDNA was amplified by PCR using the primers AAQ87743-6. The product was  
 CC cloned into the yeast expression vectors pAAH8 or pAHR8 to produce the  
 CC vectors p3A4 for the expression of the cytochrome P450 alone or p3A4R for  
 CC co-expression with the yeast NADPH-P450 reductase. The vectors are used  
 CC in a method for evaluating the safety of a chemical compound by reacting  
 CC the chemical compound with recombinantly produced human cytochrome P450  
 CC molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4,  
 CC or their auxiliary species and variants (AAQ87718-32), and yeast NADPH-  
 CC P450 reductase, either as a fused protein or in cell extracts, and  
 CC analysing the resulting metabolite to assess the safety of the chemical  
 CC compound. The method is useful for determining whether the chemical  
 CC compound, or its metabolite, will be converted into a carcinogenic or  
 CC mutagenic form through metabolism in the liver. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 CC  
 XX Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSW--ASIAFFSLLAPRRS--SLYNLQG--PNHTNY--FTGNFIDLISARTG-----E 61  
 DB 7 LAMETWLLAVSLVLLYLGTHSHGLFKLGPPTPLPFLGN---ILSYHKGFCFMDME 63

QY 62 EHAKEKYGSTLRPAGIAGAPVLNSTDPKFNHYM--KEAYDY---PKPGMAARVLRIAT 117  
 DB 64 CHKKYKVGWGF---YDG--QQPVLAITDPMIKTVLVKECYSVFTNRPPFGVGFMSKA- 117

QY 118 GDGVVTAEGEAHKKRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDMAYGES 177  
 DB 118 ---ISIAEDEWKRLRSLLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164

QY 178 AGEKKAATLET--EGVDVKDWVGRATLDVMAAGFDYKSDSLQNTKNELYVAFVGLTDGFA 236  
 DB 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNIIDSLNPNQD----- 201

QY 237 PTLDSFKAIM-WDFV--PYFRTWK--RRHEIPLTQGLAV---SRRVGIELMEQKQAVLGSA 290  
 DB 202 PFVENTKKLLRDFDLPFPFLSITVPFPIILEVLNLCVFPREVTNFRKSVKR-----M 256

QY 291 SDQAVDKKDVQGRDILSLVRANIANLPESQKLSDEEVLQISNLLPAGYETSTVLTW 350  
 DB 257 KESRLEDTKHRVDFLQMLDMSQNSKETESHKALSDELVAQSIIIFAGYETTSVLSF 316

Db	118	---	ISIAEDBEWKLRLSLLSPTTSQGLKEMVPITTAQYGDVLRNLRREA-----	164
Qy	178	AGEKATRLT-	EGVDVKDWVGRATLDVWALAGFDYKSDSLQKTNELVAFVGLTDGFA	236
Db	165	-----	ETGKPVILKDVFGAYSXDVITSTSGVNISSLNPD-----	201
Qy	237	PTLSOFKAIM-	WDFV-PYPTMK-RRHETPLTQGLAV---SRVGIHLMQKQKQAVLGS	290
Db	202	PFEVNTKLLARFDLDPFFLSITVFPFLIPLEVLNLCVPPREVTNFLRKSVKR-----	M	256
Qy	291	SDQAVDKDVGGRDILSLVRAANIAANLPESQKLSDEEVLQISNLLFAGYETSSVLTW	350	
Db	257	KESLIEDTQXRVDFQLMIDSQNSKETESHKALSDELVAQSIIIFAGYETSSVLSF	316	
Qy	351	MFHLSEDKAVQDKLRBEICQI--	DTDMPTLDEINALLPYLEAEVKESLRIDPPSPYANRE	408
Db	317	IMYELATHPVDQKQKLOBEIDAVLFPKAPPTDYTLQWEYLDWVNVETLRLFFIAMELERV	376	
Qy	409	CLKQEDFIPLAEPVIGRDSGVINEVRITKGTWVWLPLFNINRSKFYGEDAEFRPERWL	468	
Db	377	CKKQVE-----	INGMFIPLKGVVWMIYSALHRDP-KYWTPEKFLPERFS	420
Qy	469	EDVTDLSLS-IEAPYGHQASFIGSPRACFGWRFAVAEMKAFLPVTLRRVOFEP	520	
Db	421	KGNKDNIDPIYVTFPG-----	SGPRNCIGWRFALNMKALIRVLQNFSPKP	467
RESULT 7				
AAR93170				
XX	ID	AAR93170 standard; protein; 503 AA.		
XX	AC	AAR93170;		
XX	DT	11-OCT-1996 (first entry)		
XX	DE	Human cytochrome P450 molecular species 3A4 protein.		
KW	XX	Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;		
KW	KW	liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;		
KW	KW	evaluation; safety; fusion protein; metabolite; detoxification;		
XX	XX	carcinogenic.		
OS	XX	Homo sapiens.		
XX	XX	JP08056695-A.		
XX	XX	05-MAR-1996.		
XX	PF	15-JUL-1994; 94JP-00164184.		
XX	PR	20-JUL-1993; 93JP-00201120.		
XX	PR	30-JUN-1993; 93JP-00208279.		
XX	PR	17-JUN-1994; 94JP-00136053.		
XX	XX	(SUMO ) SUMITOMO CHEM CO LTD.		
XX	PA	WPI; 1996-182311/19.		
XX	DR	N-PSDB; AAT28383.		
XX	PT	Novel method for the evaluation of the safety of a cpd. - using a human		
XX	PT	cytochrome P450 and yeast NADPH reductase to determine whether the		
XX	PT	analyte cpd. is detoxified or metabolised to a carcinogen.		
XX	XX	Example 1; Page 24-26; 74pp; Japanese.		
XX	CC	This is the amino acid sequence of the human cytochrome P450 molecular		
XX	CC	species 3A4 protein. The corresp. gene was amplified from a human liver		
XX	CC	derived cDNA library as 2 fragments of 0.6 and 0.9 kb using primers		
XX	CC	AAT26933-6. The prod. was cloned into the yeast expression vector pAAH5N		
XX	CC	to generate plasmid p3A4 for prodn. of the cytochrome only or into the		
XX	CC	vector pAHRR to generate the plasmid p3A4R for co-prodn. with the yeast		
XX	CC	NADPH-P450 reductase. The sequence is placed under control of the yeast		

CC ADH gene promoter and terminator. The vectors are used in a method for  
CC evaluating the safety of a cpd. by reacting the test cpd. with  
CC recombinantly produced human cytochrome P450 mol. species 1A2 (AA28380),  
CC 2C9 (AA28381), 2E1 (AA28382), 3A4 or their variants (AA28384-98)  
CC together with yeast NADPH-P450 reductase (either as a fused protein or as  
CC a cell extract) and analysing the resultant metabolite. The cpd. is  
CC considered "safe" if it is detoxified or not rendered carcinogenic or  
CC "unsafe" if it is not detoxified or is metabolised to a carcinogenic cpd  
XX  
SQ Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;  
Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61  
DB 7 LAMETWLLLAVALSVLLVLYGTHSHGLFKLGIPOFTPLPLGN---ILSVHKGFCEFDME 63  
QY 62 EHAKEYREKYGSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYDY---PKPGMAARVLRIAT 117  
DB 64 CHKYGKVGWF---YDG--QQPVLAITDPDMIKTVLKECVSVFTNRRPFGVGFNKA- 117  
QY 118 GGVVTAEGEAHRRIRIMIPSLSAQAVKSMVPFLEKGMELVDKMDAAEKDMAVGES 177  
DB 118 ---ISIAEDEEWKRLRLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164  
QY 178 AGEKATRLT-EGVDVKDWGRATLDYMALAGFDYKSDSLQNTNELYVAFVGLTDGFA 236  
DB 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNIIDSLNPNQD----- 201  
QY 237 PTLDSFKAIM-WDFV-PYFRMK--RRHEIPTQLAV---SRRVGIELMBQKQAVLGS 290  
DB 202 PFVENTKLLRFDLDPDFLSPITVFPFLPILEVLNLCVFPREVTFNLRKSVKR-----M 256  
QY 291 SDQAVKDKVQGRDILSLVRANIANLPESQKLSDEEVLQAIISNLLFAGYETSSVLTW 350  
DB 257 KESRLDTQKRVDFLQIMIDSONSKETESHKALSDELVAQSIIFIFAGYETSSVLSF 316  
QY 351 MFHRLSEDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAFVKESLRDLPSPYANRE 408  
DB 317 IYVELATHPDVQOKLQEBIDAVLPNKAPTYDVLQMEYLDVMVNETLRFPFIAMRLERV 376  
QY 409 CLKDEDFIPLAEPIVGRDGSVINEVRIKGTVMVPLFNINRSKFIYGEDAEERPRERWL 468  
DB 377 CKKQVE-----INGMFIPKGWVMVIPSVALHRDP--KYWTEPEKELPERFS 420  
QY 469 EDVTDLSNS-IEAPYGHQASFGSPRACFGWRFVAEMKAFVLTLLRVQFEP 520  
DB 421 KKKKNIDPIYTPFG-----SGPRNCIGMRFALMNMKLLALIRVLQNFSEFKP 467

RESULT 8  
AA05202 ID AAY05202 standard; protein; 503 AA.  
XX AC AAY05202;  
XX DT 17-JUN-1999 (first entry)  
XX DE Human CYP3A4 protein.  
XX KW CYP3A4 gene polymorphism; polymorphic locus; human; altered metabolism;  
XX KW CYP3A4 substrate; drug-drug interaction identification; toxin exposure;  
XX KW genetic linkage detection; phenotypic variation.  
XX OS Homo sapiens.  
XX FN W09913106-A1.  
XX PD 18-MAR-1999.  
XX PF 02-SEP-1998; 98WO-US018158.

XX 10-SEP-1997; 97US-0058612P.  
XX (AAYS-) AAYS PHARM INC.  
XX PI Lichter JB, Guida M;  
XX WPI; 1999-215070/18.  
XX N-PSDB; AAX28295.  
XX New isolated CYP3A4 polymorphic sequences.  
XX Disclosure; Page 25-27; 40pp; English.  
CC This sequence represents the human CYP3A4 gene protein. The invention  
CC relates to a CYP3A4 sequence polymorphism, which is part of a non-  
CC naturally occurring chromosome. Nucleic acids comprising the CYP3A4  
CC polymorphic sequences can be used to screen patients for altered  
CC metabolism for CYP3A4 substrates, potential drug-drug interactions, and  
CC adverse/side effects as well as diseases that result from environmental  
CC or occupational exposure to toxins. They can also be used to establish  
CC animal, cell culture and in vitro cell-free models for drug metabolism.  
CC Polymorphic CYP3A4 gene sequences can be used for expression studies to  
CC determine the effect of promoter and/or intron sequence variations on  
CC mRNA expression and stability. The polymorphisms are also used as single  
CC nucleotide polymorphisms to detect genetic linkage to phenotypic  
CC variation in activity and expression of CYP3A4. The nucleic acids can  
CC also be used to generate genetically modified non-human animals or site  
CC specific gene modifications in cell lines  
XX Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;  
Best Local Similarity 28.7%; Pred. No. 1.3e-25;  
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61  
DB 7 LAMETWLLLAVALSVLLVLYGTHSHGLFKLGIPOFTPLPLGN---ILSVHKGFCEFDME 63  
QY 62 EHAKEYREKYGSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYDY---PKPGMAARVLRIAT 117  
DB 64 CHKYGKVGWF---YDG--QQPVLAITDPDMIKTVLKECVSVFTNRRPFGVGFNKA- 117  
QY 118 GGVVTAEGEAHRRIRIMIPSLSAQAVKSMVPFLEKGMELVDKMDAAEKDMAVGES 177  
DB 118 ---ISIAEDEEWKRLRLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164  
QY 178 AGEKATRLT-EGVDVKDWGRATLDYMALAGFDYKSDSLQNTNELYVAFVGLTDGFA 236  
DB 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNIIDSLNPNQD----- 201  
QY 237 PTLDSFKAIM-WDFV-PYFRMK--RRHEIPTQLAV---SRRVGIELMBQKQAVLGS 290  
DB 202 PFVENTKLLRFDLDPDFLSPITVFPFLPILEVLNLCVFPREVTFNLRKSVKR-----M 256  
QY 291 SDQAVKDKVQGRDILSLVRANIANLPESQKLSDEEVLQAIISNLLFAGYETSSVLTW 350  
DB 257 KESRLDTQKRVDFLQIMIDSONSKETESHKALSDELVAQSIIFIFAGYETSSVLSF 316  
QY 351 MFHRLSEDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAFVKESLRDLPSPYANRE 408  
DB 317 IYVELATHPDVQOKLQEBIDAVLPNKAPTYDVLQMEYLDVMVNETLRFPFIAMRLERV 376  
QY 409 CLKDEDFIPLAEPIVGRDGSVINEVRIKGTVMVPLFNINRSKFIYGEDAEERPRERWL 468  
DB 377 CKKQVE-----INGMFIPKGWVMVIPSVALHRDP--KYWTEPEKELPERFS 420  
QY 469 EDVTDLSNS-IEAPYGHQASFGSPRACFGWRFVAEMKAFVLTLLRVQFEP 520  
DB 421 KKKKNIDPIYTPFG-----SGPRNCIGMRFALMNMKLLALIRVLQNFSEFKP 467

```

RESULT 9
ABG68753
ID ABG68753 standard; protein; 503 AA.
XX
AC ABG68753;
XX
DT 07-OCT-2002 (first entry)
XX
DE Cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 2-13.
XX
KW Cytochrome P450; CYP3A1; CYP3A2; CYP3A3; CYP3A4; CYP3A5; CYP3A7;
XX drug metabolism; drug design; drug screening.
XX
OS Homo sapiens.
XX
PN WO200244213-A1.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-SE002631.
XX
PR 28-NOV-2000; 2000SE-00004366.
XX
PR 11-JUN-2001; 2001SE-00002661.
XX
PA (ZAPH/) ZAPHIROPOULOS P G.
XX (FINT/) FINTA C.
XX
PI Zaphiropoulos PG, Finta C;
XX
XX WPI; 2002-557532/59.
XX DR N-PSDB; ABK97692.
XX
XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
XX of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
XX metabolism, in drug design and drug screening.
XX
XX Claim 2; Fig 11; 131pp; English.
XX
XX The invention describes a cytochrome P450 protein (I) in which CYP3A43
XX exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
XX fragments, variants and multiples of (I) having essentially the same
XX characteristics. (I) is useful as a medicament, and for evaluating drug
XX metabolism, in drug design, and drug screening, and in tests for
XX adjusting the dose of drugs. This is the amino acid sequence of a novel
XX cytochrome P450 protein
XX
XX Sequence 503 AA;
XX
Query Match 13.8%; Score 394; DB 5; Length 503;
Best Local Similarity 28.7%; Pred. No. 2.1e-25;
Matches 142; Conservative 83; Mismatches 177; Indels 92; Gaps 20;
QY 47 FTGNFLDILSARTG-----EEHAKYREKYGSTLRFAGIAGAPVLNSTDPKFNHYM-XE 99
DB 46 FLGN---ILSYHKGFCMFDMCKHKKYKVGWF---YDGG--QQPVLAITPDMIKTVLVE 97
QY 100 AYDY---KPGMAARVLRIATGCVTAEGEAKHRRIMVPSLSACAVKSMVPIFLXG 156
DB 98 CYSVETNRRPFGVGFMKSA-----ISTAEDDEWRLRLSLSPFTSGKLEWVPIIAQYG 153
QY 157 MELVDKMDAAEKMDMAYGESAGEKKAATRLT-EGVDVQDVGVRATLDVNMALAGDPYKSD 215
DB 154 DVLVNLARREA-----ETGKPVTLKDVFGAYSMDVITSTSGFNWID 194
QY 216 SLQNTNELYVAVGLTGGFAPTLDSFKALM-WDFV-PYFRTMK-RRHEIPLTQGLAV-- 270
DB 195 SLANNQD-----FFVENTKGLRFDLPFFFLSTVTFPFLIPILEVINICV 240
QY 271 -SRRVGIELMEOKKQAVLGSADQAVDKVQGRDILSLVRANIANLPESOKLSDEEV 329
DB 241 PREVTNFKRSVKR-----MKESRLEDTGHRVDFLQIMDSQNSKETESHKALSDLEL 295
QY 330 LAQISNLLFAGYETSSVLTWTFHRLSEDKAVQDKLREEICQI--DTDMPTLDELNALPY 387

```

```

Db 296 VAQSIIFAGYETTSSVLSFIMYELATHPDVQOQLQBEIDAVLNLKAPPIYDTVLQMEY 355
QY 388 LEAFVXESLRLDPSPYANRECLKQEDDFIPLAEPVIGRGGSVINEVRITKGTMMMLPLFN 447
Db 356 LDMVAVNETLRLFPFIAMLRERVCKKQVE-----INGMFIKPGVGVVMPISYA 400
QY 448 INRSKFIYGEDAEPRPERLWEDVTDLSLNS-TEAPYGHQASFIISGPRACFGFRFAVAMNK 506
Db 401 LHRDP-KYWTPEKFLPERFSKKNKNDIDPIYTPFG-----SGPRNCIGNRFALMNMK 453
QY 507 AFLFVTLRRVQFEP 520
Db 454 LALIRVLQNFSPKP 467

RESULT 10
ABB08079
ID ABB08079 standard; protein; 527 AA.
XX
AC ABB08079;
XX
DT 10-SEP-2002 (first entry)
XX
DE Maize cytochrome P450, CYP72A1.
XX
KW Maize; cytochrome P450; CYP72A1; CYP92A1; plant; gene therapy; antidote.
XX
OS Zea mays.
XX
XX Key Location/Qualifiers
XX Peptide 200..527
XX FT /note= "this fragment is not provided in the protein
XX FT sequence indicated in the sequence listing"
XX PN US6380465-B1.
XX
XX 30-APR-2002.
XX
XX 11-JUL-1999; 99US-00351229.
XX
XX 12-JUL-1998; 98US-0092596P.
XX
XX (KENT ) UNIV KENTUCKY RES FOUND.
XX
XX Barrett M;
XX
XX WPI; 2002-470227/50.
XX DR N-PSDB; ABL60758.
XX
XX Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1,
XX useful for the production of plants resistant to heterocyclic,
XX PT sulfonylurea and substituted urea herbicides and organophosphate
XX insecticides e.g. Classic and Pursuit.
XX
XX Claim 13; Fig 1; 21pp; English.
XX
XX The invention relates to isolated nucleic acid molecules encoding Zea
XX mays (maize) cytochrome P450 designated CYP72A1 and CYP92A1. The
XX polypeptides can be expressed by standard recombinant methodology. The
XX nucleic acids may be used in genetic engineering protocols to transform
XX plants and other eukaryotes e.g. yeast, maize (especially), soybean,
XX CC best, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,
XX CC rose, carnation, gerbera, carrot, tomato, lettuce, chicory, pepper, melon
XX CC and cabbage. They may be used in this way to confer protection against
XX CC heterocyclic herbicides, sulfonylurea herbicides, substituted urea
XX CC herbicides and/or organophosphate insecticides. The present sequence
XX represents the Zea mays CYP72A1
XX
XX Sequence 527 AA;
XX
Query Match 13.8%; Score 392.5; DB 5; Length 527;
Best Local Similarity 24.6%; Pred. No. 3.1e-25;

```





DT 07-OCT-2002 (first entry)  
DE Cytochrome P450 CYP3A7 and CYP3A1 fusion protein.  
XX  
XX  
KW Cytochrome P450; CYP3AP1; CYP3AP2; CYP3A43; CYP3A4; CYP3A5; CYP3A7;  
KW drug metabolism; drug design; drug screening.  
XX  
XX  
OS Homo sapiens.  
XX  
XX WO200244213-A1.  
XX  
XX 06-JUN-2002.  
XX  
XX 28-NOV-2001; 2001WO-SE002631.  
XX  
XX 28-NOV-2000; 2000SE-00004366.  
XX  
XX 11-JUN-2001; 2001SE-00002061.  
XX  
XX (ZAPH/) ZAPHIROPOULOS P G.  
XX  
XX (FINT/) FINTA C.  
XX  
XX Zaphiropoulos PG, Finta C;  
XX  
XX WPI; 2002-557532/59.  
XX  
XX N-PSDB; ABK97643.  
XX  
XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets  
XX of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug  
XX metabolism, in drug design and drug screening.  
XX  
XX Claim 1; Page 73-75; 131pp; English.  
XX  
XX The invention describes a cytochrome P450 protein (I) in which CYP3A43  
XX exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub  
XX fragments, variants and multiples of (I) having essentially the same  
XX characteristics (I) is useful as a medicament, and for evaluating drug  
XX metabolism, in drug design, and drug screening, and in tests for  
XX adjusting the dose of drugs. This is the amino acid sequence of a novel  
XX cytochrome P450 protein  
XX  
XX Sequence 535 AA;  
XX  
XX  
XX Query Match 13.5%; Score 385.5; DB 5; Length 535;  
XX Best Local Similarity 23.0%; Pred. No. 1.3e-24;  
XX Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;  
XX  
XX 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61  
XX  
XX 7 LAVETWLLAVSLILLVLYGTRTHGLFKLGIPOGTPPLPGLN---ALSPKGYWTFDME 63  
XX  
XX 62 BHAKYREKYGSTLRFAGIAGAPVLSNDPKVFNHVM-KEAYD-----YPRKGMAARVL 113  
XX  
XX 64 CYKRYKRWG-----YDCCQPMIAITDPMIKTVLKECYSVFTNRRPFGPVGFMKNAI 118  
XX  
XX 114 RIATGDGVVTAEGEAKHRRHIMIPSLSAQAVKSMVPIFEKGMELVDKQMDAEKDWA 173  
XX  
XX 119 SI-----AEDEENKIRLSLPTTSGLKKEWPIIAQYGDVLRNLRREA----- 164  
XX  
XX 174 VGESAGEKKAATLET-EGVDYKDWGGRATLDVMAAGFDYKSLQNKTNELVAVGLT 232  
XX  
XX 165 -----ETGKEPVLKRVFGAYSDVITSTSGVSDLSLNNPDQ-----PVENT 207  
XX  
XX 233 D---GFAPLDSFKAIMNDVFPVPTMKRRHEIPLTQGLAVS---RRVGELMEQKKQAV 286  
XX  
XX 208 KXLLRFPN-LDPFVLISIKVF-PFL-----TILEALNITVPKVISFLKSVKQIK 257  
XX  
XX 287 LGSASDAQVDKXVQGR--DILSLVRANIANIIPESQKLSDBEVLQAQINLLPAGYETS 344  
XX  
XX 258 EGRL-----KETQKRVDFLQMLDMSQSKDSETHKALSDELMAQSIIFAGYETT 310  
XX  
XX 345 STVLTMFHRLESQKAVODKLRREICQTDQM-----PTLDELNALPYLEAFVKESIRLD 399  
XX  
XX 311 SSVLSFIITELIATHPDVQOKVQKE---IDTVLPKAPPTDVLQLELYDMVNVNITLRF 367

QY 400 PPSPYANRECLKDEDFIPLAEPVIGRSGSVINEVITKGTWMLPLFNINRSKFIYGEDA 459  
DB 368 FVAMLERVCKKQVE-----INGMFIPKGVVWIPSYVLHDP-KYWTETP 411  
QY 460 EBFRRPERWLEDVTDLSNS-IEAPYGHQASFIISGRACFGWFAVAENKAFILFVTLRRVQF 518  
DB 412 EKFLPERFSKKNKNDIPVITPFG-----SGPRNCIGMRFALVNMKMLALVRVLQNF 465  
QY 519 EP 520  
DB 466 KP 467  
RESULT 13  
AAE26192  
ID AAE26192 standard; protein; 502 AA.  
XX  
XX AAE26192;  
XX  
XX 14-NOV-2002 (first entry)  
XX Human cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) protein.  
XX Human; cytochrome P450; subfamily IIIA; polypeptide 5 isogene; CYP3A5;  
XX drug screening; polymorphism; haplotype; drug metabolising disorder;  
XX gene therapy.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 30 /note= "This amino acid changes to Tyr due to single  
XX nucleotide polymorphism"  
XX Misc-difference 100 /note= "This amino acid changes to Tyr due to single  
XX nucleotide polymorphism"  
XX  
XX WO200246209-A2.  
XX  
XX 13-JUN-2002.  
XX  
XX 07-DEC-2001; 2001WO-US047218.  
XX  
XX 08-DEC-2000; 2000US-0254367P.  
XX 03-MAY-2001; 2001US-0288470P.  
XX (GENA-) GENAISSANCE PHARM INC.  
XX Anastasio AE, Han J, Klieem SE, Rounds E;  
XX WPI: 2002-636448/68.  
XX N-PSDB; AAD43350, AAD43351.  
XX Novel isolated polynucleotide which is a polymorphic variant of  
XX cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) gene useful for  
XX expressing CYP3A5 protein isoform used in drug screening techniques.  
XX  
XX Claim 29; Fig 3; 127pp; English.  
XX  
XX The invention relates to isolated polynucleotide having cytochrome P450,  
XX subfamily IIIA, polypeptide 5 isogene (CYP3A5). The invention is useful  
XX for screening drugs. The invention is useful for studying expression and  
XX function of CYP3A5 and expressing CYP3A5 protein for use in screening for  
XX candidate drugs to treat diseases related to CYP3A5 activity. The  
XX polymorphism and haplotype data is useful for validating whether CYP3A5  
XX is a suitable target for drugs to treat drug metabolising disorders,  
XX screening for such drugs and reducing bias in clinical trials of such  
XX drugs. The invention is also useful for therapeutic purposes. The  
XX invention is useful in studying the effect of variation on the biological  
XX activity of CYP3A5 as well as on the binding affinity of candidate drugs  
XX to CYP3A5, or for studying the enzymatic properties of such CYP3A5  
XX variants using these candidate drugs as substrate. The invention is

CC useful in gene therapy. The present sequence is human CYP3A5 protein

XX

SQ Sequence 502 AA;

Query Match 13.5%; Score 385; DB 5; Length 502;

Best Local Similarity 28.1%; Pred. No. 1.3e-24;

Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAFSW--ASIAFFSLYL-APRRSLYNLQ-PNHTNY-FTGNFLDILSARTG-----E 61

DB 7 LAVETWLLAVSLVLLYLYGTRTHGLFRKLGIPGPTPLPLGN---VLSYRQGLWKFDTE 63

QY 62 EHAKYREKYGSLTRPAGIAGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATG-- 118

DB 64 CYKKGKMGWT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111

QY 119 ----DOVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDQAV 174

DB 112 GFWMKSAISLAEDDEWKIRSLSPFTTSGKLEWFPPIIAQYGDVLRNLARE-AEKG--- 167

QY 175 GESAGEKKATRLTEGVDVQKMGVGRATLDVNALAGFDYKSDSLON-----KTNELYVAF 228

DB 168 -----KPVTLKIDFGAYSMDVITGTSFGVNIDSLNPNQDPFVSTKELKF 213

QY 229 VGLTDFGAPTLDSFKAIMWDFVPYFRTMKRRHEIPLTQGLAVS--RRVGIELMEQKQAV 286

DB 214 ----GFLDPL-FLSILFPFL-----TFVFEALNVSLFPKDTINFLSKS----- 252

QY 287 LGSASDAQVDKQVQGRDILSLVLRANLANLPESQKLSDEEVLQAINLIFAGYETSST 346

DB 253 VNRMKKSLRNDKQKRLDFLQIMIDSQNSKETESHKALSDELAQAQSIIFIFAGYETSS 312

QY 347 VLTWPHRLSDEKAVQDKLREIEICQI--DTDMPTLDLNLALPYLEAFVKESLRDPPSPY 404

DB 313 VLSFTLYELATHPDVQOKLQKSIDAVLPKAPPTYDVAVQMEYLDVMVNETLRFPVAVR 372

QY 405 ANRECLKQSDFTPLAEFVIGRQGVINEVRIKGTWMLPLFNINRSKFIYGEDAEFRP 464

DB 373 LERTCKKQVE-----INGVFIPKSMVVIPTVYALHDP-KYWTPESEFRP 416

QY 465 ERWLEDVDTSLNS-IEAPYGHQASFTSGPRACFGRFAVABMKAPLFTVLRVQFEP 520

DB 417 ERFSSK-KUSIDPIYITPFG-----TGRNCIGRFPALNMKALIRVLQNFSEFKP 466

RESULT 14

ID ADE57186

AC ADE57186 standard; protein; 502 AA.

XX

XX ADE57186;

XX

XX 29-JAN-2004 (first entry)

XX

XX Human Protein P20815, SEQ ID NO 3047.

XX

XX DE

XX

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX

XX OS Homo sapiens.

XX

XX PN WO2003016475-A2.

XX

XX PD 27-FEB-2003.

XX

XX PF 14-AUG-2002; 2002WO-US025765.

XX

XX PR 14-AUG-2001; 2001US-0312147P.

XX

XX PR 01-NOV-2001; 2001US-0346382P.

XX

XX PR 26-NOV-2001; 2001US-0333347P.

XX

XX (GEO ) GEN HOSPITAL CORP.

XX (FARB ) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

XX GENBANK; P20815.

XX

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX

XX Claim 1; Page; 1017pp; English.

XX

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 502 AA;

Query Match 13.5%; Score 385; DB 7; Length 502;

Best Local Similarity 28.1%; Pred. No. 1.3e-24;

Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAFSW--ASIAFFSLYL-APRRSLYNLQ-PNHTNY-FTGNFLDILSARTG-----E 61

DB 7 LAVETWLLAVSLVLLYLYGTRTHGLFRKLGIPGPTPLPLGN---VLSYRQGLWKFDTE 63

QY 62 EHAKYREKYGSLTRPAGIAGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATG-- 118

DB 64 CYKKGKMGWT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111

QY 119 ----DOVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDQAV 174

DB 112 GFWMKSAISLAEDDEWKIRSLSPFTTSGKLEWFPPIIAQYGDVLRNLARE-AEKG--- 167

QY 175 GESAGEKKATRLTEGVDVQKMGVGRATLDVNALAGFDYKSDSLON-----KTNELYVAF 228

DB 168 -----KPVTLKIDFGAYSMDVITGTSFGVNIDSLNPNQDPFVSTKELKF 213

QY 229 VGLTDFGAPTLDSFKAIMWDFVPYFRTMKRRHEIPLTQGLAVS--RRVGIELMEQKQAV 286

DB 214 ----GFLDPL-FLSILFPFL-----TFVFEALNVSLFPKDTINFLSKS----- 252

QY 287 LGSASDAQVDKQVQGRDILSLVLRANLANLPESQKLSDEEVLQAINLIFAGYETSST 346

DB 253 VNRMKKSLRNDKQKRLDFLQIMIDSQNSKETESHKALSDELAQAQSIIFIFAGYETSS 312

QY 347 VLTWPHRLSDEKAVQDKLREIEICQI--DTDMPTLDLNLALPYLEAFVKESLRDPPSPY 404

DB 313 VLSFTLYELATHPDVQOKLQKSIDAVLPKAPPTYDVAVQMEYLDVMVNETLRFPVAVR 372



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17.5 Seconds  
(without alignments)  
1643.181 Million cell updates/sec

Title: US-10-066-007-3  
Perfect score: 2852  
Sequence: 1 MFILVLTGALGAARSWAS.....RIVGREKGYQMRLOVKPVE 557

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	4	US-09-518-386B-1
2	2852	100.0	557	4	US-09-518-386B-3
3	396.5	13.9	503	4	US-09-144-357-2
4	370	13.0	504	1	US-08-457-274A-25
5	370	13.0	504	5	PCT-US95-05758-25
6	367	12.9	520	4	US-09-527-073-2
7	345	12.1	524	4	US-09-976-594-533
8	338.5	11.9	503	4	US-09-583-447A-2
9	331	11.6	540	4	US-09-302-620B-99
10	330.5	11.6	507	1	US-08-457-274A-22
11	330.5	11.6	507	5	PCT-US95-05758-22
12	328	11.5	504	4	US-09-583-447A-4
13	328	11.5	540	4	US-09-302-620B-98
14	326.5	11.4	541	3	US-09-158-767-19
15	326.5	11.4	541	3	US-09-158-767-20
16	321	11.3	576	3	US-08-948-564-16
17	318.5	11.2	526	1	US-08-298-426-4
18	294.5	10.3	489	4	US-08-852-087-4
19	294	10.3	522	4	US-09-302-620B-97
20	289	10.1	522	4	US-09-302-620B-96
21	276	9.7	420	4	US-09-583-447A-6
22	276	9.7	467	4	US-09-126-420A-17
23	269.5	9.4	517	4	US-09-302-620B-100
24	269.5	9.4	517	4	US-09-911-781-32
25	269	9.4	507	1	US-08-457-274A-23
26	269	9.4	507	5	PCT-US95-05758-23
27	267.5	9.4	517	4	US-09-302-620B-101

28	263	9.2	523	4	US-09-302-620B-95
29	259.5	9.1	510	3	US-08-948-564-4
30	259	9.1	510	4	US-09-852-067-2
31	251.5	8.8	508	4	US-09-126-420A-25
32	246	8.6	512	4	US-09-302-620B-103
33	243	8.5	512	4	US-09-302-620B-102
34	241.5	8.5	476	1	US-08-313-075A-30
35	236.5	8.3	498	1	US-08-457-274A-24
36	236.5	8.3	498	5	PCT-US95-05758-24
37	235.5	8.3	499	4	US-09-302-620B-104
38	235.5	8.3	504	4	US-09-126-420A-18
39	233.5	8.2	509	4	US-09-499-302A-8
40	228	8.0	426	2	US-08-560-398-4
41	227	8.0	504	4	US-09-499-302A-6
42	226.5	7.9	500	3	US-09-292-768-68
43	226.5	7.9	500	3	US-09-292-768-70
44	225.5	7.9	500	3	US-09-292-768-4
45	223.5	7.8	524	4	US-09-126-420A-24

ALIGNMENTS

RESULT 1  
US-09-518-386B-1  
; Sequence 1, Application US/09518386B  
; Patent No. 6365386  
; GENERAL INFORMATION:  
; APPLICANT: HOSHINO, Tatsuo  
; APPLICANT: OJIMA, Kazuyuki  
; APPLICANT: SETOGUCHI, Yutaka  
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
; CURRENT APPLICATION NUMBER: US/09/518,386B  
; CURRENT FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: EP 99104668.1  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EP 00101666.6  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 557  
; TYPE: PRT  
; ORGANISM: Phaffia rhodozyma  
; NAME/KEY: TRANSIT  
; LOCATION: (1)..(26)  
US-09-518-386B-1

Query Match	100.0%;	Score 2852;	DB 4;	Length 557;
Best Local Similarity	100.0%;	Pred. No. 4.4e-278;	Mismatches 0;	Indels 0;
Matches 557;	Conservative 0;			Gaps 0;
QY	1	MFILVLTGALGAARSWASIAFFSLYLA	PRSSLYNQLGPNHTNYFTGNFLDILSARTG	60
Db	1	MFILVLTGALGAARSWASIAFFSLYLA	PRSSLYNQLGPNHTNYFTGNFLDILSARTG	60
QY	61	BEHAKYREKYGSTLRPAGIAGAVLNSTDP	KVFNHVKAYDPKPGMAARVLRATG	120
Db	61	BEHAKYREKYGSTLRPAGIAGAVLNSTDP	KVFNHVKAYDPKPGMAARVLRATG	120
QY	121	VVTEGEAHKRRHRIMIPSLSAQAVKSMVPI	FLFKGMELVDKMDAEAKDMAVGESAGE	180
Db	121	VVTEGEAHKRRHRIMIPSLSAQAVKSMVPI	FLFKGMELVDKMDAEAKDMAVGESAGE	180
QY	181	KKATRLTEGVVDKDWGRATLDVMA	LADYKSDSLQNTNELYYVAVFLGTDG	240
Db	181	KKATRLTEGVVDKDWGRATLDVMA	LADYKSDSLQNTNELYYVAVFLGTDG	240
QY	241	SFKAIMDFFPYFRMKRRHEIPLTQGLAV	SRVVGIELMEQKQAVLGSASQAVDKDV	300
Db	241	SFKAIMDFFPYFRMKRRHEIPLTQGLAV	SRVVGIELMEQKQAVLGSASQAVDKDV	300

QY 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWTFHRLSEDKA 360  
 DB 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWTFHRLSEDKA 360  
 QY 361 VODKLRREICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDDEDFIPLAE 420  
 DB 361 VODKLRREICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDDEDFIPLAE 420  
 QY 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFRPERWLEDTVDSLSNIEA 480  
 DB 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFRPERWLEDTVDSLSNIEA 480  
 QY 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540  
 DB 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540  
 QY 541 GREKEGYQMRLOVKPVE 557  
 DB 541 GREKEGYQMRLOVKPVE 557

RESULT 2  
 US-09-518-386B-3  
 ; Sequence 3, Application US/09518386B  
 ; Patent No. 6365386  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOSHINO, Tatsuo  
 ; APPLICANT: OJIMA, Kazuyuki  
 ; APPLICANT: SETOGUCHI, Yutaka  
 ; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
 ; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
 ; CURRENT APPLICATION NUMBER: US/09/518,386B  
 ; CURRENT FILING DATE: 2000-03-03  
 ; PRIOR APPLICATION NUMBER: EP 99104668.1  
 ; PRIOR FILING DATE: 1999-03-09  
 ; PRIOR APPLICATION NUMBER: EP 00101666.6  
 ; PRIOR FILING DATE: 2000-02-01  
 ; NUMBER OF SEQ ID NOS: 32  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 557  
 ; TYPE: PRT  
 ; ORGANISM: Phaffia rhodozyma  
 US-09-518-386B-3

Query Match 100.0%; Score 2852; DB 4; Length 557;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-278;  
 Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFILVLLTGALGAAFSWASIAFFSLYLAPRSSLYNLOGPNHTNYFTGNFLDILSARTG 60  
 DB 1 MFILVLLTGALGAAFSWASIAFFSLYLAPRSSLYNLOGPNHTNYFTGNFLDILSARTG 60  
 QY 61 EEHAKYREKYGSTLRAGIAGAPVLSNTPKVFNHVMEAYDYPKPGMAARVLRATGDG 120  
 DB 61 EEHAKYREKYGSTLRAGIAGAPVLSNTPKVFNHVMEAYDYPKPGMAARVLRATGDG 120  
 QY 121 VVTAEGSAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDROMEDAAEKOMAVGESAGE 180  
 DB 121 VVTAEGSAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDROMEDAAEKOMAVGESAGE 180  
 QY 181 KKAATRLTEGVDDKVMVGRATLDVMAAGFDVKSDSLQNKTNELVYVAVGLTDGFPATLD 240  
 DB 181 KKAATRLTEGVDDKVMVGRATLDVMAAGFDVKSDSLQNKTNELVYVAVGLTDGFPATLD 240  
 QY 241 SFKAIMWDFPYFRTMKRREIIFLTQGLAVSRVAGIELMEQKKQAVLGSASDAQVKKDV 300  
 DB 241 SFKAIMWDFPYFRTMKRREIIFLTQGLAVSRVAGIELMEQKKQAVLGSASDAQVKKDV 300  
 QY 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWTFHRLSEDKA 360  
 DB 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWTFHRLSEDKA 360

QY 361 VODKLRREICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDDEDFIPLAE 420  
 DB 361 VODKLRREICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDDEDFIPLAE 420  
 QY 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFRPERWLEDTVDSLSNIEA 480  
 DB 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFRPERWLEDTVDSLSNIEA 480  
 QY 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540  
 DB 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540  
 QY 541 GREKEGYQMRLOVKPVE 557  
 DB 541 GREKEGYQMRLOVKPVE 557

RESULT 3  
 US-09-144-367-2  
 ; Sequence 2, Application US/09144367  
 ; Patent No. 6432639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lichter, Jay  
 ; APPLICANT: Guido, Marco  
 ; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4  
 ; FILE REFERENCE: SEQ-12P  
 ; CURRENT APPLICATION NUMBER: US/09/144,367  
 ; CURRENT FILING DATE: 1998-08-31  
 ; PRIOR APPLICATION NUMBER: 60/058,612  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 503  
 ; TYPE: PRT  
 ; ORGANISM: H. sapiens  
 US-09-144-367-2

Query Match 13.9%; Score 396.5; DB 4; Length 503;  
 Best Local Similarity 28.7%; Pred. No. 5.8e-31;  
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSW--ASTAFFSLYLAPRRS--SLYNLOG--PNHTNY--FTGNFLDILSARTG-----E 61  
 DB 7 LAMETWLLLVALLVLLYLYGTHSHGLPKKLGIPQPTPLPLGN---ILSYHKGFCHMEDNE 63  
 QY 62 EHAKYREKYGSTLRAGIAGAPVLSNTPKVFNHVME--KEAYDY---PKPGMAARVLRAT 117  
 DB 64 CHKYGKVMGF---YDG--QQPVLAITDPMIKTVLKECYSVFTNRRRPFPGVGFMKSA- 117  
 QY 118 GGVVTAEGSAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKWMEDAAEKDMAVGES 177  
 DB 118 ---ISIASDEEWKRLRSLSPFTSGKLKEMVPIAQYGDVLRNLREA----- 184  
 QY 178 AGEKKAATLET--EGVDVDDKVMVGRATLDVMAAGFDVKSDSLQNKTNELVYVAVGLTDGFA 236  
 DB 165 -----ETGKPVTLKDVFGAYSDVITSTSGVNIIDSLNNPD----- 201  
 QY 237 PILDSEFKAIM--WDFV--PYFRTMK--RRHEIPLTQGLAV---SRRVGIELMEQKKCAVLGSA 290  
 DB 202 PFVETNKKLLRFDLDPDFLSITVFPFLIPLEVLNLCVFPFVNFLRKSVRK-----M 256  
 QY 291 SDQAVDKDQVQGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTW 350  
 DB 257 KESRIEDTCQHRVDFQLQMLDSQNSKETESHKALSDLELVAQSIIFIFAGYEITSSVLSF 316  
 QY 351 MFHRLSEDKAVQDKUREIICQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408  
 DB 317 INVETLATHEDVQOKLQEEBIDAVLPENKAPPTTYDTVLQMEYLDVMVNETLRLFFIAMLERY 376  
 QY 409 CLKDEDFIPLAEFVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEFRPERWL 468

```

Db 377 CKDVE-----INGMEIPKGVWVMPISYALHRDP-KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-LEAPYGHQASFGSPRACFGWPAFAEMKAFVTLRRVQPEP 520
Db 421 KKNKNDIDPIYTPFG-----SGPRNCIGMRPALMMKALIRVLQNFSPK 467

RESULT 4
US-08-457-274A-25
; Sequence 25, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,274A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat
; STRAIN:
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
US-08-457-274A-25

Query Match 13.0%; Score 370; DB 1; Length 504;
Best Local Similarity 26.2%; Pred. No. 2.7e-28;
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY 7 LTGALGLAASFWSAIFSLY-LAPRSSLYNLQG-----PNHTNYFTGNF-LDI 54
Db 3 LLSALTLETWLLAWVLVLYGFGTRTHGLFKQGPCKPLPFFGVLNLYMGLWKFDV 62
QY 55 LSARTGEHAKYREKYGSTLRFAGIAPVLNSTDPKFNHVM-KEAY-----DYPKP 106
Db 63 -----ECHKKYKIGW---LFDG--QMLPFAITDTEMIKNVLKBCFSVFTNRDRFGV 111
QY 107 GMAARVLRIATGDGVVTAEGSAHKRRHRIIMPISLSAQAVKSMVPIFLEKGMELVDROMED 166
Db 112 GI-----MGKAVSVAKDEWKRYALLSPFTTSGLKXEMPIIEQYGDILVYKQK 163
QY 167 AABKDMAYGESAGEKKAATRLT-EGVDVKWVGRATLDVNALAGFDYKSDSLQN----- 219

```

```

Db 164 A-----ETGKEVTMKKVFSGYSMDVITSTSGVNVDSLNNPKDPFV 204
QY 220 -XTNEL-----YVAFVGLTDFAPTLDSFKAIMW--DFVPYFRTMKRHEIPL 264
Db 205 EKTCKLLRDFDPPLFLSVLFPFLT-----PIYEMLNICMFPKDSIEFFK----- 250
QY 265 TQGLAVSRVGIELMEQKQAVLGASDAQVDKDVQGRDILSLVLRANIAANLPESQ-K 323
Db 251 -----KFYRMKETRLDSVQKHRV-----DFQLMMNAHNSDKKESHTA 290
QY 324 LSDDEVLAQISNLLPAGYETSTVLTMMFHLRSEDKAVQDKLREIEICOI---DTDMPTLDE 381
Db 291 LSDMEITAQSIIFIPAGYETSTVLTMMFHLRSEDKAVQDKLREIEICOI---DTDMPTLDE 350
QY 382 LNALPYLEAFVKESLRDLPPSPYANRECKLQEDFPLAEPVIGRDSVINEVRITKGMV 441
Db 351 VMEMEYLDVNLNETLRLYPIGNRLERLVCKKQVE-----INGVFWPKGSVV 395
QY 442 MLPLFNINRSKFTYGEDAEFFPERKLEVDVTDLSNS-LEAPYGHQASFGSPRACFGWRF 500
Db 396 MIPSYALHRDPQHWPE-PPEFRPERFSKENKGSIDPYVILPFG-----NGPRNCIGNRF 448
QY 501 AVAEMKAFVTLRRVQPEP 520
Db 449 ALMMYKALTKVLQNFSPQ 468

RESULT 5
PCT-US95-05758-25
; Sequence 25, Application PC/TUS9505758
; GENERAL INFORMATION:
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05758
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat
; STRAIN: Unknown
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:

```

CHROMOSOME/SEGMENT: Unknown  
PCT-US95-05758-25

Query Match 13.0%; Score 370; DB 5; Length 504;  
Best Local Similarity 26.2%; Pred. No. 2.7e-28;  
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY 7 LTGALGAFAFASWASIAFFSLY-LARRSSLYNLOG-----PNTNYFTGTF-LDI 54  
DB 3 LLSALTLETWLLAVLVLYGFGTHTGLFKKQIPGPKPLPFGFVLYNYMGLWIFDV 62  
QY 55 LSARTGEEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHV-KEAY-----DYPKP 106  
DB 63 -----ECHKKYKING--LEFG-QMPLFAITDEMIKNVLKCEFSVFTNRDGPV 111  
QY 107 GMAARVLRATGDVVTAAGBAHRRIMISLSAQAVKSVPIFLKGMELVDKWMED 166  
DB 112 GI-----MGKAVSVADEEKWRALLSTFTSGRLKEMFPIIEQYGDILVKYKQE 163  
QY 167 AAEKDMAVGESAGEKATRLT-EGVDVKDWVGRATLDVMAAGDPYKSDSLON----- 219  
DB 164 A-----ETGKPVTKKVFAGYAMDVITSTSGVNVDSLNNPKDPFV 204  
QY 220 -KNEL-----YVAVGLTGDGAPTLDSFKALMW--DVPYFRTMKRRHSIPL 264  
DB 205 ETKKLLRFDFFPLFLSVLPFFLT-----PIYEMLNICMPKDSIEFFK----- 250  
QY 265 TQGLAVSRVGIEMLEKQKQAVLGSASQAVDKQVQGRDILSLVRANIANLPESQ-K 323  
DB 251 -----KFVYRMKETRLDSVQHRV-----DFQLMMNAHNSDKESHSTA 290  
QY 324 LSDEEVLAQISNLLFAGYTSSTVLTWFMHRLSEKAVQDKUREIICQI--DTMPILDE 381  
DB 291 LSDMEITAQSIIFAGYBPTSTSLSVLHSLATHPTDKKQLQEEIDRALPNKAPTYDT 350  
QY 382 LNALPYLEAFVKESLRDPPSPYANRECKDDEFTPLAEPVIGRDSGVINEVRITKGMV 441  
DB 351 VMEVEYLDWNLNTLLYFIGNRLERVCKDVE-----INGVFPKGSVV 395  
QY 442 MLPLFNIRSKFYGBDAEERPERWLEDVTDLSNS-TEAPYGHQASISGPRACFGWRF 500  
DB 396 MIPSYALHRDPQHWPE-PSEFRPERFSKENKSIDPVYVLPFG-----NGFRNCIGMRF 448  
QY 501 AVAEMKAFVTLRRVQFEP 520  
DB 449 ALMMKALTKVLQNFSPQ 468

RESULT 6  
US-09-527-073-2  
; Sequence 2, Application US/09527073  
; Patent No. 6534313  
; GENERAL INFORMATION:  
; APPLICANT: Michael M. Neff  
; APPLICANT: Joanne Chory  
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING  
; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING  
; FILE REFERENCE: SALKINS.024  
; CURRENT APPLICATION NUMBER: US/09/527,073  
; CURRENT FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: US 60/124570  
; PRIOR FILING DATE: 1999-03-16  
; PRIOR APPLICATION NUMBER: US 60/170,931  
; PRIOR FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: US 60/172,832  
; PRIOR FILING DATE: 1999-12-20  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 520  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana

US-09-527-073-2

Query Match 12.9%; Score 367; DB 4; Length 520;  
Best Local Similarity 22.2%; Pred. No. 5.7e-28;  
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;  
QY 4 LVLTGALGAFAFASWASIAFFSLYLAARR-----SSLYNLOGPNHNTNYFTGTFDILS--- 56  
DB 13 VLVLSVLISLVKGMSL-----LWRRPKITEEHFSKQIRGPPY-HFFIGNVKELVGMML 67  
QY 57 -----ARTGEEHAKYREKYGSTLRFAGIAGAPVLNST--DPKVFNHVWKEA 100  
DB 68 KASSHPMPFSNINLPRVLSFYHWRKIYGATFL--VWFGTFTLTVADDLIREFESKS 124  
QY 101 YDYPGPGMAARVLRATGDVVTAAGBAHRRIMIPSLSAQAVKSMVPIFLKGMELV 160  
DB 125 BFEYK-NEAHLPLVKQLEGDLGLSLKGEKWAHRRKIISPTFHMENKLLVDPVLKSVTDMV 183  
QY 161 DQWEDAAEKDMAVGESAGEKATRLTETEGVDVKDWVGRATLDVMAAGDPYKSDSLQNK 220  
DB 184 DKMSDKLSN-----GEVE-----VDVYENFQIILTEDVISTAFGSSYE----- 222  
QY 221 TNELYVAVFGLTDGFAP-----TLDSFKALMWDFVPYFRTMKRRHEIFLTQGLAV 270  
DB 223 -----DGEAVFRLQAQQMLLCAEAFQV--FIPGYRFFP-----TRGNLX 260  
QY 271 SRVVG-----IEMEQKQAVLGSASQAVDKQVQGRDILSLVRANIANLPESQK 323  
DB 261 SKLDEKEIRKSLKLIERRQNAIDGEGEC--KEPAKDLGLGLMICA-----KN 308  
QY 324 LSDEEVLAQISNLLFAGYTSSTVLTWFMHRLSEKAVQDKUREIICQI--DTMPILDE 381  
DB 309 VTVDIVIECKSFFPAGKQTTSNLLTWTLLSMHPESQAKARDEVLRVCGSRDVPKDH 368  
QY 382 LNALPYLEAFVKESLRDPPSPYANRECKDDEFTPLAEPVIGRDSGVINEVRITKGMV 441  
DB 369 VYVLTSLMILNESLRLYPPIVATIRRAKSDVK-----LGGYKIPQGTTEL 413  
QY 442 MLPLFNIRSKFYGBDAEERPERWLEDVTDLSNS-TEAPYGHQASIS--GPRACFGW 498  
DB 414 LPIITAVHDDQAINWDVNEFNPAFADGVPRAAK-----HPVGFIPGLGVRTICIG 466  
QY 499 RFPAEMKAFVTLRRVQFEPFISHPEYEHITLII 534  
DB 467 NLAILQAKLTAVMIQRTFTH--LAPTYQHAPTVL 499

Query Match 12.1%; Score 345; DB 4; Length 524;  
Best Local Similarity 23.4%; Pred. No. 9.4e-26;  
Matches 140; Conservative 104; Mismatches 207; Indels 148; Gaps 25;

US-09-976-594-533  
; Sequence 533, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 533  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 6825202CD1  
; US-09-976-594-533



QY 2 FIVLLTALGALAA--FSWASIAFPFSLYLAAPRSSLNLOGPNHTNYFTGNFLLDILSART 59  
 DB 19 WLLLLVGGSWLLARVLAWT-----YTFYDNCRR--LQCFPPQPKQNWFWGHOGLVPTPTEE 72  
 QY 60 GEEH-----AKYREKY-----GSTLRFAGIAGAPVLNSTDPKVFNVHMKEA-----100  
 DB 73 GMKTLTQLVVTPYPOGFKLWLGPTF-----PLLLCHPDIIIRPITSASAAVAPKDMIF 124  
 QY 101 YDYKPGVAARVLIATGDGVVTVABGAHGRHRIMIPSLSAQAVKSMVPIFLEKGMVELV 160  
 DB 125 YGFLKPMV-----GDGLLSGGKGRHRRMLTTPAFHFNILKPYMKIP--NKSVNI--173  
 QY 161 DKMMEDAAEKDMAVGESAGEKATRLTEG---VDVKDWVGRATLDVMALAGDYKSDSL 217  
 DB 174 ---YHD-----KWQRLASEGSRALDMFEHISLWLSLQKCVSPFES--NC 214  
 QY 218 QNKTNELVAVFGTGDGPAFLDSFKAIMDVPY-----PR-----TWK 257  
 DB 215 QKPTSE--YIAAILSEAFVEXRNOQIILLHTDLYLTFDQGRFRACHLVHDFDVAIOE 273  
 QY 258 RHHEIPLTQGLAVSRVVGIELMEOKKQAVLGSASDAQVKKQVQGRDILSLVRANIAAN 317  
 DB 274 RRTULP--TQGD-----DFKNAKSKTLDFTDVLKSKDEGKE-----312  
 QY 318 LPESQKLSDEEVLQISNLLPAGYETSTVLTWMFHLSEDKAVQDKLREEICQIDTDM 377  
 DB 313 -----LSDEDIRAEADTFMEGHDTTASGLSWLYHLAKHPYEQECROEVQELLKDR 366  
 QY 378 TL-----DELNALPYLEAEVKESLDPSPVYANRECLKDEDFPLAEPVIGRDGVSINEV 433  
 DB 367 PIEIEWDQAOLPFTMCKESLRLHPVPVVISRCCTQD-----FVLPDGRV-----413  
 QY 434 RITKTMVLMPLFNINRSKFIYGDABEFPRERWLEDVTDLSNLSIEAPYGHQASFTSGPR 493  
 DB 414 -IPKGVCLINIGIHYNPTVM--PDPEVYDPFRE-----DQENIKERSPLAFTPFSGAPR 466  
 QY 494 ACFGRPVAEMKALFVTLRRVQPEPIISHPEVHEHTLISRPVIGREKEGYQWRLQ 552  
 DB 467 NCIGQAFVAMKVVLTLLHFRILPHTHEPR-----RKPELILRAEGGLWLRVE 517

RESULT 8  
 US-09-583-447A-2  
 ; Sequence 2, Application US/09583447A  
 ; Patent No. 6645745  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WOJNOWSKI, Leszek  
 ; APPLICANT: GELLNER, Klaus  
 ; APPLICANT: EISELT, Regina  
 ; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A  
 ; TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX  
 ; FILE REFERENCE: 310115.401  
 ; CURRENT APPLICATION NUMBER: US/09/583,447A  
 ; CURRENT FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 503  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-583-447A-2

Query Match 11.9%; Score 338.5; DB 4; Length 503;  
 Best Local Similarity 27.1%; Pred. No. 4e-25;  
 Matches 128; Conservative 77; Mismatches 187; Indels 81; Gaps 17;  
 QY 61 EEHAKYREKYSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYDYKPKQMAARVLIATGD 119  
 DB 63 ECNEKYGMWG---LYEG--QQPMLVIMDPDMIKTLVKECYSVFTNQMPGPFGLK-S 116  
 QY 120 GVVTAEGEAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDMAVGESAG 179

DB 117 ALSFAEDEEWKIRITLLSPAFTSVKFKEMVPIISQGDMLVRSRLQEAEB-----165  
 QY 180 EKXATRLTEGVQDVQKDWGVRATLDVMALAGPDYKSDSLQNKTNELVAVFVGLTDGFAPTL 239  
 DB 166 -----NSKSINLKDOFFGAYTMDVITGTLFGVNLDSLNNPD-----PFL 204  
 QY 240 DSPKAIM--WDFVPYFRITMKRHEIPLTOGLA--VSRRVGIELMEOKKQAVLGSASQAVDK 297  
 DB 205 KMKKLLKLDLDFDPFL-----ISLFPFLTPVFEALNIGLFPKDVTHFLKNSIERMKES 259  
 QY 298 --KDVQR--DIISLIVRANIAANLPSOKLSDEEVLQISNLLFAGYETSSTVLTWMEH 353  
 DB 260 RLKDKQKHVRDFFQOMIDNSQSKSHKALSDLELVAOSIIIIIFAAVDTTSTLFFIMY 319  
 QY 354 RLSEDKAVQDKLREEICQIDTDM-----TLDELNALPYLEAEVKESLRLDPPSPYANRE 408  
 DB 320 ELATHPDVQKLOEE---IDAVLPKAPVTYDALVQMEYLDVMVNETLRLFPVWSKRVTRV 376  
 QY 409 CLKDEDPIABPVIGRDGVSINEVITKGTWMLPLFNINRSKFIYGEDABEFPERWL 468  
 DB 377 CKKDIE-----INGVFIKGLAMVFIYALHHPD--KYWTEPEKFCPERFS 420  
 QY 469 EDVTDLSNLSIE--APYGHQASFTSGPRACFGWRPFAVEMKALFVTLRRVQPEP 520  
 DB 421 KKNKDSIDLRYIPFG-----AGPRNCIGMRPALTNIKLAIVIRALQNFSEXP 467

RESULT 9  
 US-09-302-620B-99  
 ; Sequence 99, Application US/09302620B  
 ; Patent No. 6331420  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wilson, C. Ron  
 ; APPLICANT: Craft, David L.  
 ; APPLICANT: Eirich, Dudley  
 ; APPLICANT: Eshoo, Mark  
 ; APPLICANT: Madduri, Krishna M.  
 ; APPLICANT: Cornett, Cathy A.  
 ; APPLICANT: Brenner, Alfred A.  
 ; APPLICANT: Tang, Maria  
 ; APPLICANT: Loper, John C.  
 ; APPLICANT: Gleeson, Martin  
 ; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450  
 ; TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA  
 ; TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS  
 ; TITLE OF INVENTION: RELATING THERETO  
 ; FILE REFERENCE: 1010-16.seq  
 ; CURRENT APPLICATION NUMBER: US/09/302,620B  
 ; CURRENT FILING DATE: 1999-04-30  
 ; NUMBER OF SEQ ID NOS: 109  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 99  
 ; LENGTH: 540  
 ; TYPE: PRT  
 ; ORGANISM: Candida tropicalis  
 US-09-302-620B-99

Query Match 11.6%; Score 331; DB 4; Length 540;  
 Best Local Similarity 23.9%; Pred. No. 2.6e-24;  
 Matches 135; Conservative 94; Mismatches 205; Indels 132; Gaps 22;  
 QY 42 NHTNYF-----TGNLDILSARTGEE--HAK-----YREKYG 71  
 DB 25 NYTRWYVPIPLVLLSLNFIHLTKYLERFRHAKLGNVVDFTFGIATPLILILKSKG 84  
 QY 72 STLRAFA-----GIAGAPVLNSTDPKVFNVHMKEAAYDYKPKQMAAR 111  
 DB 85 TWKPFAMSWNNKIYKDPKKTGLRIVGLPLEITDPENIKAVLATQNFDSLGRHD 144  
 QY 112 VLRITATGQVVTAEGEAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKD 171  
 DB 145 FLYSLGDIETLGDAGWKHSRTMLRPQFAREQVSHV-----KLEPHVQVF 191

QY 172 MAVGESAGEKATRLTEGVYKDVWGRATLDVMAAGFDYKSDSLQNKTNELVAVFL 231  
Db 192 F-----KVRKRGQTFDQELFFRLTVDSATEFLFGSAESLRDSS-----VGL 236  
QY 232 TDGFAPTLDSFKAIMWDFVYFERTK--RRHEIPITQGLAVSRVRGIELMEQKQAVLGS 289  
Db 237 T-----FTTKDFEG-RGDFADAFNSQYQAVRFLQQYWLIN--GAFF--RKSIIVHK 287  
QY 290 ASQAVDK-----KDVQGRDILSLVRANIANLPESQKLSDEVLQ-QISNLLFAGYE 342  
Db 288 FADHYVQKALELTDLDDQLQKQDGVFEL-----AKQTRDPKVLDRDQLNLLVAGRD 339  
QY 343 TSSTVLTWMPHRLSDEKAVQDKLREI-----CQIDTDMPTLDLALNLPYLAFAFKES 395  
Db 340 TTAGLSFVFFYELSRNPEVFAKLREENVNRFGLGEEARVEEISFESLAKSCYLAKEVINEA 399  
QY 396 LRLDPPSPVANRECLKDEDFIPLAEVPIGRDGSVINEVRITKGTWMLPLFNINRSKFIY 455  
Db 400 LRLYPSVPHNFRVATRNTT-LPRGG---GKDG--CSP1VVKKGQVVMYTVIGTHRDPFIY 453  
QY 456 GEDAEFRPERWLEDVTDLSNLSIEAPYGHQASFGSPRACFGWRFAVAKMAFLFVTLRR 515  
Db 454 GADADVFRPERWEPETRKLWAYVP-----FNGGPRICLGCQFALTEAS---YTVRL 504  
QY 516 VQ-----PEPIISHPEYEHITLIS 535  
Db 505 LQEFNLSLDPNAEYPPKLQNTLTL 530

## RESULT 10

US-08-457-274A-22  
; Sequence 22, Application US/08457274A  
; Patent No. 5734086  
; GENERAL INFORMATION:  
; APPLICANT: Scott, Jeffrey G.  
; APPLICANT: Tomita, Takashi  
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Nixon, Hargrave, Devans & Doyle  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 716-263-1304  
; TELEFAX: 716-263-1600  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 507 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Musca domestica  
; STRAIN: Rutgers

; DEVELOPMENTAL STAGE: Adult  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: Chromosome 5  
; US-08-457-274A-22  
Query Match 11.6%; Score 330.5; DB 1; Length 507;  
Best Local Similarity 23.5%; Pred. No. 2.6e-24;  
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;  
QY 10 ALGLAFAWASIAFSLYLAPRRSSLYNLOGPNH--TNYFTGNFLDLSA-RIGEEHAKY 66  
Db 10 ALGV-----LASLALYFVRNFGYWKRGIPHEEPHLVMGNVKGSRKYHIGELIADY 62  
QY 67 RKYGSTLRFAGI--AGAPVLNSTDPKVFNVHM--KEAYDYPKPGMAARVLRIATGQVVT 123  
Db 63 YRKFKGSGPFGAGIFLGHKPAAVLDKELRKRVLIKDFSNFANRGLYYNEXDDPLTGHVLM 122  
QY 124 ARGEAHKRRHRIIMPSLSAQAVKSMVFIFLEKGMELVDKMDAAAEKDMVAGESAGEKKA 183  
Db 123 VEGERSLRKLSTFTTAGKMYNTVLEVGORLLEVME----- 164  
QY 184 TELETG--VDVKDVGGRATLDVMAAGFDYKSDSLQ-----KTNELYV 226  
Db 165 -KLEVSSSLDMRDILARENTDVGSAFGIECSLNRPHDRFLAMGRKSTEVPRHNALIM 223  
QY 227 AFVGLTDGFAPTLDSFKAIMWDFVYFERTKRRHEIPLTQGLAVSRVRGIELMEQKQAV 286  
Db 224 AFI-----DSFPEL-----SRKLGMRVLPEVDVHQF 248  
QY 287 LGSASDAQVD--KKDVQGRDILSLVRANIANLPESQK-----LSDEEVLQATISNLLFA 339  
Db 249 FMSSIKETVDYREKKNIRNDFLDLVDLK--NNPESISKGLGLTFNELAAQVVFVFLG 305  
QY 340 GYETSTVLTWMPHRLSDEKAVQDKLREI-----CQIDTDMPTLDLALNLPYLAFAFKES 395  
Db 306 GFETSSSTMGFALYELAQNLQDLREENVNEVDFQFEDNISYDALMNPYLDQVLNET 365  
QY 396 LRLDPPSPVANRECLKDEDFIPLAEVPIGRDGSVINEVRITKGTWMLPLFNINRSKFIY 455  
Db 366 LRKYPVGSALTQTLNDY-----VYPHPKYV-----LPKGTLVIPVLGIHYDPELY 413  
QY 456 GEDAEFRPERWLEDVTDLSNLSIEAPYGHQASFGSPRACFGWRFAVAKMAFLFVTLRR 515  
Db 414 -PNPEFDPERESPVMKQSDVD-----WLGFGDGPGRNCIGMRFKQVQSRGLGLALVIRH 467  
QY 516 VOFE-----PIISHPE 526  
Db 468 FRFTVCSRTDIPMQINPE 485

## RESULT 11

PCT-US95-05758-22  
; Sequence 22, Application PC/TUS9505758  
; GENERAL INFORMATION:  
; APPLICANT: Cornell Research Foundation, Inc.  
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05758  
; FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1304  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 507 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
STRAIN: Rutgers  
DEVELOPMENTAL STAGE: Adult  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Chromosome 5  
PCT-US95-05758-22

Query Match 11.6%; Score 330.5; DB 5; Length 507;  
Best Local Similarity 23.5%; Pred. No. 2.6e-24;  
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;  
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;

QY 10 ALGUAFAWASIAFSSLYLAPRSSLNLOQPH--TNVFTGNFLDLSA-RTGEHAKY 66  
DB 10 ALGV-----LASLALFYFWMNFYWKRRGIPHEEPHLVMGVKSLRKYHIGIADY 62  
QY 67 REKVGSTLRPAGI--AGAPVLNSTDPKVFNVHM--KEAYDYPKPGMAARVLRIATGDGVVT 123  
DB 63 YRKFGSGPAGIFLGHKPAAVLDKELRXLVIKDFSNFANRGLYNEKDDPLATGLVM 122  
QY 124 AEGFAHRHRRIMTIPSLSAQAVKSMVPIFLEKGMELVDKMDAEDAAEKMDAVGESAGEKAT 184  
DB 123 VEGEKWRSRLTKLSPTFTAGMKMTYNTVLEVGQRLLEVME-----KTNELVY 226  
QY 184 TRLETEG--VDYKDWVGRATLDVMAAGFDYKSDSLON-----LSDDEVLAQISNLLFA 339  
DB 165 -KLEVSSELDNRDILARNFTDVGSAFGEICSLNRPDHRFLAMGRKSEIVPRNHALIM 223  
QY 227 AFVGLTDFAPTLDSFKAIMWDFVYPFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAV 286  
DB 224 AFI-----DSFPEL-----SRKLGRLVPEDVHQF 248  
QY 287 LGSASDAQVD---KDVQGRDILSLYRANIAANLPESQK-----LSDEVLAQISNLLFA 339  
DB 249 FMSSTIKETVDYREKNNIRNDFDLVLDLK---NNPESISKGLGLTFNEALAAQVFFFLG 305  
QY 340 GYETSSTVLTWVHRLSEDKAVQKLEEEI---CQIDTDMPTLDELNALPYLEAFVKES 395  
DB 306 GFETSSTWGTALYELAQNLQQLDRLEEVNEVDFQKEDNISYDALMNPYLDQVLNET 365  
QY 336 LRLDPPSYANRECKDEDFIPLAEVPIGRDGSVINEVRITKGTMMWMLPLFNINRSKFIY 455  
DB 366 LRKYPVGSALTQTLNDY-----VYPHPKV-----LPKGTLPFIPVLGHIHPDELY 413  
QY 456 GDAEERPEPEWLEVDVTLNSIEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRR 515  
DB 414 -NPPEFDFEPESEVMVQKRSVD-----WLFGDGRNCGNIGRFGQWQSRLGALVIRH 467  
QY 516 VQFE-----PIISHPE 526  
DB 468 PRFTVCSRTDIPMQINPE 485

RESULT 12  
US-09-583-447A-4  
; Sequence 4, Application US/09583447A

Patent No. 6645745  
GENERAL INFORMATION:  
APPLICANT: WOUNOWSKI, Leszek  
APPLICANT: GELLNER, Klaus  
APPLICANT: EISELT, Regina  
TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A  
TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX  
FILE REFERENCE: 310115.401  
CURRENT APPLICATION NUMBER: US/09/583,447A  
CURRENT FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 504  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-583-447A-4

Query Match 11.5%; Score 328; DB 4; Length 504;  
Best Local Similarity 27.1%; Pred. No. 4.6e-24;  
Matches 127; Conservative 75; Mismatches 185; Indels 82; Gaps 18;  
Matches 127; Conservative 75; Mismatches 185; Indels 82; Gaps 18;

QY 68 EYKVGSTLRPAGI--AGAPVLNSTDPKVFNVHM--KEAYDYPKPGMAARVLRIATGDGVVT 124  
DB 66 ERYG---EMWGLYEGQQPMLVMDPMIKTVLKECYSVFTNQPLGPMGLK-SALSFA 121  
QY 125 EGEAHRHRRIMTIPSLSAQAVKSMVPIFLEKGMELVDKMDAEDAAEKMDAVGESAGEKAT 184  
DB 122 EDEWKRIITLSPAPTSVKFKEMVPIISQCCDMLVRSLOEAE----- 165  
QY 185 RUTEQGVVDKDWGRATLDVMAAGFDYKSDSLONKTNELVYVAFVGLTDGFPATLDSFKA 244  
DB 166 --NSKINILKDFGAVTMDVITGTLFGVNLDSLNNPQ-----PFLKNMK 209  
QY 245 IM-WDEVPVFRMTKRHEIPLTQGLA--VSRVGIEMLEKQKQAVLGSASDAQVDK--KDV 300  
DB 210 LKLLDPLDFELL-----ISLFFPLTPVFEALNIGLFPKDVTHFLKNSIERMKESRLKDK 264  
QY 301 QGR--DILSLVPRANTANLPESQKLSDEVLAOISNLLFAGYETSSTVLTWVHRLSED 358  
DB 265 QHRVDFVQQMDISQNSKETSKSHKALSDLELVAQSIIFAIAYDTTSTLPIYVELATH 324  
QY 359 KAVQDKLRBEICQIDTDMPT-----TIDELNALPYLEAFVKESRLDPPSPYANRECKDE 413  
DB 325 PDVQKQLEH---IDAVLPNKAPVTVDALVQNEYLDMVNETLRLFPVVSRTVCKKDI 381  
QY 414 DTIPLAEVPIGRDGSVINEVRITKGTMMWMLPLFNINRSKFIYGEDAEERPE-RWLEDTV 472  
DB 382 E-----INGVFIPKGLAVMVPYIALHDDP-KYWTPEFKPCPSRFSKKNK 425  
QY 473 DSLNSIE-APYGHQASFTSGPRACFGWRFAVAEMKAFVTLRRVQFEP 520  
DB 426 DSIDLRYTFFG-----AGPRNCIGMRALTNIKLAIVIRALQNFSPKP 468

RESULT 13  
US-09-302-620B-98  
; Sequence 98, Application US/09302620B  
; Patent No. 6331420  
GENERAL INFORMATION:  
APPLICANT: Wilson, C. Ron  
APPLICANT: Craft, David L.  
APPLICANT: Birch, Dudley  
APPLICANT: Behoo, Mark  
APPLICANT: Madduri, Krishna M.  
APPLICANT: Cornett, Cathy A.  
APPLICANT: Brenner, Alfred A.  
APPLICANT: Tang, Maria  
APPLICANT: Loper, John C.  
APPLICANT: Gleeson, Martin  
TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450  
TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA  
TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS

```

; TITLE OF INVENTION: RELATING THERETO
; FILE REFERENCE: 1010-16.sq
; CURRENT APPLICATION NUMBER: US/09/302,620B
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Candida tropicalis
; US-09-302-620B-98

```

```

Query Match 11.5%; Score 328; DB 4; Length 540;
Best Local Similarity 23.9%; Pred. No. 5.1e-24;
Matches 133; Conservative 93; Mismatches 197; Indels 134; Gaps 22;

QY 42 NHTNYP-----TGNFDILSARTGEE--HAK-----YREKYG 71
DB 25 NYTRWYFYFPLVLLSLNFISSLHTRYLERRRFAKPLGNFVRDPTFGIATPLLLIYLSKG 84
QY 72 STURFA-----GIAGAPVLNSTDPKVENHVMKEAYDYPKPGMAAR 111
DB 85 TVNKPFWGLWNNKYIVDRPKYKTTGLRIVGLPLIETMPENIKAVLATQFNDFSLGTRHD 144
QY 112 VLIATGDDGVVTAEGBAHGRHRIMIPSLSAQAVKSMVPIFEKGMELVDQMDEAAEKD 171
DB 145 FLYSLLDGDIPTLDGAGWKHSRTMLRPOPAREQVSHV-----KLLBPHVQVF 191
QY 172 MAVGESAGKATRLTEGVGVKMDVGRATLDVMALAGFDYKSDSLQNKTNELVAFVGL 231
DB 192 F-----XHVKKHGGQTFDQELFFRLTVDSATEFLFGSABSLRDES-----IGL 236
QY 232 TDGFATLDSFKAMWDFVYFRTWK--RRHEIPTQGLAVSRVRVIGIELMEQKQAVLGS 289
DB 237 T-----PTTKDFDC--RRDFADAENYSQYQAYRFLIQMYWILN--GSEF--RKSIAVYHK 287
QY 290 ASDQAVDK-----KDVQGRDILSLIVRANIANLPESQKLSDEEVLQ--QISNLLFAGYE 342
DB 288 FADHYVQKALELTDQLQDQGVFLYEL-----AKQTRDPKVRDQLNLVLVGRD 339
QY 343 TSSTVLTMFHLRSEDKAVQDKLREFI-----CQIDTDMPTLDELNALPYLEAFVKES 395
DB 340 TTAGLLSFVYELSRNPEVFAKLREVENRFGLBSEARVEEISFESLASCSEYLVKAVINET 399
QY 396 LRDPSPYANRECLKDEDFIPLAEVPIGDSVINEVRITKGTVMPLPLNINRSKFIY 455
DB 400 LRLYPSVPHNFRVATNTT--LPRGG--GEDG--YSPIVVKGVQVMVTVIATHRDPISY 453
QY 456 GEDAEPRPRERWLEDVTDLSNLSIAPYGHQAFISGPRACFGWRFAVAKMKAFLFVTLRR 515
DB 454 GADADVPRPRERWEPETRKLWAYVP-----FNGGPRICLQGFALTEAS---YVTVRL 504
QY 516 VQPEPIISHPEYDHITL 532
DB 505 LQ-----EFAHLSM 513

```

```

RESULT 14
US-09-158-767-19
; Sequence 19, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: FR 97-12094
; EARLIER FILING DATE: 1997-09-24

```

```

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Altered sequences
; US-09-158-767-19

Query Match 11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.6%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

QY 78 GIA--GAPVLNSTDPKVENHVMKEAYD--YKPGMAARVLRIATGDDGVVTAEGBAHGRHR 134
DB 73 GVARRGGVLVTTCDFRNLHVLKARFDNYPKPGFWHGFERDLGLDGI FNSDGDGTWLAQRK 132
QY 135 IMIPLSAQAVKSMVPIFEKGMELVDKMDAAEKDMAVGESAGKATRLTEGVGVK 194
DB 133 TAALFETTRTURTAMSRWVSRSH--GRLLPILA--DAAKGKQAQ-----VDLQ 176
QY 195 DWVGRATLDVMALAGFDYKSDSLQ--KTNELVAFVGLTDGFPATLDSFKAIMWDFVPY 252
DB 177 DLLRLTLTFDNICGLAFGKDPETLAQGLPENEFASAFDTEA--TLNRF-----IFPE 227
QY 253 F-----RTWKRHEIPTQGLA--VSRVIGIELMEQKQAVLGSASDOAVDKDVQGRDIL 306
DB 228 FLWRCKKWLGLMETTTTSSMAHYDQYLAAVIKRKLLEAAGNGKCDTAAATHD-----DLL 283
QY 307 SLLVRANIANLPESQKLSDEEVLQISNLLFAGYETSTVLTMMFHLRSEDKAVQDKLR 366
DB 284 SPMFKG-----SYSDSLQHVAFNLGRDITSSVALSWFFWLVTSTHPEVERKIV 334
QY 367 EIIQI-----DMDMETLDELNALPYLEAFVKESLRDPPSPYANRECLKDE 413
DB 335 RELCSVLAASRGADHPALMLAEPTFELDRVLVYKALSETLRLYPSVPEDSKHVVD-- 393
QY 414 DIPLAEVPIGRDGSVINEVRITKGTVMPLPLNINRSKFIYGEDAEFEPRERWLE---- 469
DB 394 DVLV-----DGTFF-----VPAGSSVTVSYISAGRMKGVWGEDCLFEPRWLSADGT 440
QY 470 --DVTDSNLSIAPYGHQAFISGPRACFGWRFAVAKMKAFLFVTLRVQFEPILSHPEY 527
DB 441 KPEQDHSYKVV-----AFNAGPRVCLGKDLAYLQMKNIAGSVLLRHLRTVAPGRHVE 492
QY 528 EHITLISRPRIVGREKEGYQWRLOVKEP 555
DB 493 QKMSLTFM-----KGG--LRMEVRP 511

RESULT 15
US-09-158-767-20
; Sequence 20, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: FR 97-12094
; EARLIER FILING DATE: 1997-09-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Altered sequences
US-09-158-767-20

Query Match      11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.8%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

Qy 78 GIA--GAPVLNSTDPKVFNHVMKEAYD-YPKPGMAARVLRIRATGQVVTABGEAHKRHR 134
Db 73 GVARRGLVTVTCDPRNLEHVKARFDNYKGFVHGVFDLLGDI FNSDGDITWLAQRK 132
Qy 135 IMPSLSAQAVKSMVPIFLKGMELVDKMDAEDKMDAVGESAGEKAKATRLTEGVYK 194
Db 133 TAALEFTRTLRTAMGRWSRSIH--GRLLPILA--DAAKGKAAQ-----VDLQ 176
Qy 195 DWVGRATLDVMALAGFDYKSDSLQN--XTNELYVAFVGLTDGFAPTLDSFKAIMWDFVPY 252
Db 177 DLLRLTFDNLICGLAFGKDEPTELAQGLPENEFASAFDRATEA---TLNRF-----IFPE 227
Qy 253 F-----RTMKRRHEIPLTQGLA-VSRRVGIELMEQKQAVLGSASDAQVDKXDVQGRDIL 306
Db 228 FLWRCKKWLGLGMEHTTLTSSMAHVQDYLAAVIKKRKLELAAGNGKCDTAATHD----DLL 283
Qy 307 SLLVRANIAANLPESQKLSDEEVLQAIISNLPAGYETSTVLTMMFHRLSDEKAVQDKLR 366
Db 284 SRFRKRG-----SYDESIOHVALNFILAGRTSSVALSWFPLVSTHFAVERKIV 334
Qy 367 EEICQI-----DTMPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDE 413
Db 335 RELCSVLAASRGADHPALWLAEPPTFEELDRVLVLAALSETLRLYSPVSPEDSKHVAD- 393
Qy 414 DFIPILAEFVIGRDSVINEVRIITKGTWMLPLENINRSKIYGEDAEERPERWLE---- 469
Db 394 DYLP-----DGTG-----VPAGSSVTYIISAGRMKGWGEDCLEFRPERWLSADGT 440
Qy 470 --DVTDSLNSIEAPYGHQASFGISGFACFGWRFAVEMKAFLEVTLLRVQFPIISHPEY 527
Db 441 KFEQHDSYKFV-----AFNAGPRVCLGKDLAYLQMKNIAGSVLLRHRLTVAPGHRVE 492
Qy 528 EHITLISRPRIVGREKEGYQMRLOQXP 555
Db 493 QKMSLTLEF-----XGG--LRMEVRP 511

Search completed: April 2, 2004, 14:02:16
Job time : 19.5 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 14:00:49 ; Search time 40 Seconds  
(without alignments)  
3653.561 Million cell updates/sec

Title: US-10-066-007-3  
Perfect score: 2852  
Sequence: 1 MFILVLTGALGAAPSWAS.....RIVGREKEGYQRLQKPV 557

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1070241 seqs, 262374223 residues

Total number of hits satisfying chosen parameters: 1070241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pap.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT NEW PUB.pap.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW PUB.pap.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW PUB.pap.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW PUB.pap.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pap.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pap.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pap.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pap.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pap.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pap.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pap.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW PUB.pap.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW PUB.pap.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	14	US-10-066-007-1
2	2852	100.0	557	14	US-10-066-007-3
3	398.5	14.0	503	15	US-10-313-963A-56
4	396.5	13.9	503	14	US-10-146-575-2
5	385.5	13.5	503	9	US-09-957-997-3
6	384	13.5	537	12	US-10-425-114-38180
7	377.5	13.2	527	12	US-10-425-114-67055
8	376.5	13.2	547	12	US-10-425-114-65616
9	368.5	12.9	547	12	US-10-425-114-61218
10	368.5	12.9	560	12	US-10-425-114-59349
11	368.5	12.9	562	12	US-10-425-114-59350
12	367	12.9	520	9	US-09-992-901-2
13	365.5	12.8	620	12	US-10-425-114-38827
14	365	12.8	520	12	US-10-114-270-60
15	353	12.4	524	12	US-10-424-599-199559

16	352	12.3	533	12	US-10-424-599-269344	Sequence 269344, App
17	350	12.3	520	14	US-10-072-012-548	Sequence 548, App
18	350	12.3	520	14	US-10-032-189-115	Sequence 115, App
19	349.5	12.3	520	12	US-10-424-599-228819	Sequence 228819, App
20	349	12.2	508	14	US-10-274-694-16	Sequence 16, Appl
21	348.5	12.2	546	15	US-10-291-365-405	Sequence 405, App
22	347.5	12.2	524	15	US-10-291-465-877	Sequence 877, App
23	347.5	12.2	524	15	US-10-291-465-878	Sequence 878, App
24	347.5	12.2	524	15	US-10-291-265-879	Sequence 879, App
25	346.5	12.1	524	9	US-09-989-722-264	Sequence 264, App
26	346.5	12.1	524	9	US-09-989-723-264	Sequence 264, App
27	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
28	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
29	346.5	12.1	524	9	US-09-989-731-264	Sequence 264, App
30	346.5	12.1	524	9	US-09-989-732-264	Sequence 264, App
31	346.5	12.1	524	9	US-09-991-073-264	Sequence 264, App
32	346.5	12.1	524	9	US-09-990-442-264	Sequence 264, App
33	346.5	12.1	524	9	US-09-991-163-264	Sequence 264, App
34	346.5	12.1	524	9	US-09-993-604-264	Sequence 264, App
35	346.5	12.1	524	9	US-09-990-456-264	Sequence 264, App
36	346.5	12.1	524	9	US-09-989-721-264	Sequence 264, App
37	346.5	12.1	524	9	US-09-992-598-264	Sequence 264, App
38	346.5	12.1	524	9	US-09-989-293A-264	Sequence 264, App
39	346.5	12.1	524	9	US-09-989-735-264	Sequence 264, App
40	346.5	12.1	524	9	US-09-990-444-264	Sequence 264, App
41	346.5	12.1	524	9	US-09-991-181-264	Sequence 264, App
42	346.5	12.1	524	9	US-09-989-730-264	Sequence 264, App
43	346.5	12.1	524	9	US-09-990-438-264	Sequence 264, App
44	346.5	12.1	524	9	US-09-993-687-264	Sequence 264, App
45	346.5	12.1	524	10	US-09-989-734-264	Sequence 264, App

## ALIGNMENTS

### RESULT 1

US-10-066-007-1  
; Sequence 1, Application US/10066007  
; Publication No. US2003007691A1  
; GENERAL INFORMATION:  
; APPLICANT: HOSHINO, Tatsuo  
; APPLICANT: OJIMA, Kazuyuki  
; APPLICANT: SETOGUCHI, Yutaka  
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
; CURRENT APPLICATION NUMBER: US/10/066,007  
; CURRENT FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: US/09/518,386  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: EP 99104668.1  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EP 00101666.6  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 557  
; TYPE: PRT  
; ORGANISM: Phaffia rhodozyma  
; FEATURE:  
; NAME/KEY: TRANSIT  
; LOCATION: (1)...(26)  
US-10-066-007-1

Query Match 100.0%; Score 2852; DB 14; Length 557;  
Best Local Similarity 100.0%; Pred. No. 2.3e-248;  
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFILVLTGALGAAPSWASIAFFSLVLAAPRSSLYNLOQPNHNTYFTGNFDILSARTG 60

Db 1 MFILVLTGALGAAPSWASIAFFSLVLAAPRSSLYNLOQPNHNTYFTGNFDILSARTG 60

QY 61 BEHAKYREKYGSLRFRFAGIAGAPVLNSTDPKVFNVHNVKGAIDYIPKPGMAARVLRIATG 120

```
Db 61 EEHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNHVKAEADYDPKPGMAARVLRATG 120
Qy 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMVAGESAGE 180
Db 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMVAGESAGE 180
Qy 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELYNLYVAFVGLTDCGFAPTLD 240
Db 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELYNLYVAFVGLTDCGFAPTLD 240
Qy 241 SFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAVLGASDAQVKKDV 300
Db 241 SFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAVLGASDAQVKKDV 300
Qy 241 SFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAVLGASDAQVKKDV 300
Db 241 SFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAVLGASDAQVKKDV 300
Qy 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
Db 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
Qy 361 VQDKLREEICQIDTDMPTLDLDELNALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLDELNALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
Qy 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEPRPRWLEDTVDSLSNIEA 480
Db 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEPRPRWLEDTVDSLSNIEA 480
Qy 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
Db 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
Qy 541 GREKEGYQMRLOQKPE 557
Db 541 GREKEGYQMRLOQKPE 557
```

## RESULT 2

US-10-066-007-3

; Sequence 3, Application US/10066007

; Publication No. US20030077691A1

; GENERAL INFORMATION:

; APPLICANT: HOSHINO, Tatsuo

; APPLICANT: OJIMA, Kazuyuki

; APPLICANT: SETOGUCHI, Yutaka

; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE

; FILE REFERENCE: ASTAXANTHIN SYNTHETASE

; CURRENT APPLICATION NUMBER: US/10/066,007

; CURRENT FILING DATE: 2001-02-01

; PRIOR APPLICATION NUMBER: US/09/518,386

; PRIOR FILING DATE: 2000-03-03

; PRIOR APPLICATION NUMBER: EP 99104668.1

; PRIOR FILING DATE: 1999-03-09

; PRIOR APPLICATION NUMBER: EP 00101666.6

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: Patent in ver. 2.1

; SEQ ID NO 3

; TYPE: PRT

; LENGTH: 557

; ORGANISM: Phaffia rhodozyma

US-10-066-007-3

```
Query Match 100.0%; Score 2852; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.3e-248;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MFILVLITGALGAALFASWASIAFFSLYLAPRRSSLYNLOGPNHTNYFTGNFLDILSARTG 60
Db 1 MFILVLITGALGAALFASWASIAFFSLYLAPRRSSLYNLOGPNHTNYFTGNFLDILSARTG 60
Qy 61 EEHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNHVKAEADYDPKPGMAARVLRATG 120
Db 61 EEHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNHVKAEADYDPKPGMAARVLRATG 120
```

```
Qy 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMVAGESAGE 180
Db 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMVAGESAGE 180
Qy 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELYNLYVAFVGLTDCGFAPTLD 240
Db 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELYNLYVAFVGLTDCGFAPTLD 240
Qy 241 SFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAVLGASDAQVKKDV 300
Db 241 SFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVGIEMLEKQKQAVLGASDAQVKKDV 300
Qy 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
Db 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
Qy 361 VQDKLREEICQIDTDMPTLDLDELNALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLDELNALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
Qy 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEPRPRWLEDTVDSLSNIEA 480
Db 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEPRPRWLEDTVDSLSNIEA 480
Qy 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
Db 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
Qy 541 GREKEGYQMRLOQKPE 557
Db 541 GREKEGYQMRLOQKPE 557
```

## RESULT 3

US-10-313-963A-56

; Sequence 56, Application US/10313963A

; Publication No. US20040002078A1

; GENERAL INFORMATION:

; APPLICANT: Boutell, Jonathan

; APPLICANT: Godber, Benjamin

; APPLICANT: Hart, Darren

; APPLICANT: Blackburn, Jonathan

; TITLE OF INVENTION: Arrays

; FILE REFERENCE: KIL-001

; CURRENT APPLICATION NUMBER: US/10/313,963A

; CURRENT FILING DATE: 2003-06-19

; PRIOR APPLICATION NUMBER: US 60/335,806

; PRIOR FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: US 60/410,815

; PRIOR FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 56

; LENGTH: 503

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-313-963A-56

```
Query Match 14.0%; Score 398.5; DB 15; Length 503;
Best Local Similarity 28.7%; Pred. No. 6.2e-27;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
```

```
Qy 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLOG--PNHTNY--FTGNFLDILSARTG-----E 61
Db 7 LMETWLLVALLSLVLLYLYGTHSHCLPKKLIGPPTPLFLGN---ILSYHKGFCMEDME 63
Qy 62 EHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNHVM--KEADY---PKPGMAARVLRAT 117
Db 64 CHKKYKVGWGF---YDG--QQPVLAITDPDMIKTVLKECYSVFTNRRPFGPVGFMKSA- 117
Qy 118 GDCVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMVAGES 177
```

```

Db 118 ---ISIAEDDEWKLRLSLLSPFTSCKLKEMVPIIAQYGDVLRNLRREA----- 164
Qy 178 AGEKATRLLET-EGVDVKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAVGLTDGFA 236
Db 165 -----ETGKPVTLKDVFGAYSDVITSTSGVNIIDSLNNPQD----- 201
Qy 237 PTLDSFKAIM-WDFV-PYFRMK-RRHEIPTQGLAV---SRRVGIELMEQKQAVLGSA 290
Db 202 PFVENTKLLRFDFLDPFFLSIVFPFLPILEVLNICVFPREVTFNLRKSVKR-----M 256
Qy 291 SDQAVDKOVQGRDILSLVRANIAPESQKLSDEEVLAQISNLLFAGYETSTVLTW 350
Db 257 KESRLDTQKRVDFLQMLDMSQSKETESHKALSDELVQAQSIIFIFAGYETTSVLSF 316
Qy 351 MEHRLSEDKAVQKLRREICQI--DTPMTLDDELNALPYLEAFVKESLRLLDPPSPYANRE 408
Db 317 IMVELATHPDVQKQLEIDAVLPNKAPPTDYVQLQMEYLDVMVNETLLEFPFIAMLERLV 376
Qy 409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRWL 468
Db 377 CKKQVE-----INGMFIPKGVVWVIMPSYALHRDP-KYWTEPEKFLPERFS 420
Qy 469 EDVTDLSNS-IEAPYGHQASFIISGPRACFGHFAVAEMKAEFLVTLRRVQREP 520
Db 421 KKNKDNIDPYIYTPFG-----SGPRNCIGMRFPALMMKMLALIRVLQNFSPRP 467

RESULT 4
US-10-146-575-2
; Sequence 2, Application US/10146575
; Publication No. US200030059800A1
; GENERAL INFORMATION:
; APPLICANT: Guido, Marco
; APPLICANT: Lichter, Jay
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-122
; CURRENT APPLICATION NUMBER: US/10/146,575
; PRIOR FILING DATE: 2002-05-14
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 503
; TYPE: PRI
; ORGANISM: H. sapiens
US-10-146-575-2

Query Match
Best Local Similarity 28.7%; Score 396.5; DB 14; Length 503;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

Qy 13 LAAPSW--ASTAFSLYIAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLAVSLVLLLYGTHSHGLFKLGIPTLPFLGN---ILSYHKGFCMFDMVE 63
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLSNSTDPKVFNHVM-KEAYD---PKPGMAARVLRIAT 117
Db 64 CHKKYKGVWG---YDG---QQPVLAITDPMIKTVLKECYSVFTNRRPFGPVGFMKSA- 117
Qy 118 GDGVVTAEGEAHRRHRRIMPSLSAQAVKSMVPIFLKGMELVDKMMEDAAEKDMAVGES 177
Db 118 ---ISIAEDDEWKLRLSLLSPFTSCKLKEMVPIIAQYGDVLRNLRREA----- 164
Qy 178 AGEKATRLLET-EGVDVKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAVGLTDGFA 236
Db 165 -----ETGKPVTLKDVFGAYSDVITSTSGVNIIDSLNNPQD----- 201
Qy 237 PTLDSFKAIM-WDFV-PYFRMK-RRHEIPTQGLAV---SRRVGIELMEQKQAVLGSA 290
Db 202 PFVENTKLLRFDFLDPFFLSIVFPFLPILEVLNICVFPREVTFNLRKSVKR-----M 256
Qy 291 SDQAVDKOVQGRDILSLVRANIAPESQKLSDEEVLAQISNLLFAGYETSTVLTW 350
Db 257 KESRLDTQKRVDFLQMLDMSQSKETESHKALSDELVQAQSIIFIFAGYETTSVLSF 316
Qy 351 MEHRLSEDKAVQKLRREICQI--DTPMTLDDELNALPYLEAFVKESLRLLDPPSPYANRE 408
Db 317 IMVELATHPDVQKQLEIDAVLPNKAPPTDYVQLQMEYLDVMVNETLLEFPFIAMLERLV 376
Qy 409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRWL 468
Db 377 CKKQVE-----INGMFIPKGVVWVIMPSYALHRDP-KYWTEPEKFLPERFS 420
Qy 469 EDVTDLSNS-IEAPYGHQASFIISGPRACFGHFAVAEMKAEFLVTLRRVQREP 520
Db 421 KKNKDNIDPYIYTPFG-----SGPRNCIGMRFPALMMKMLALIRVLQNFSPRP 467

```

```

Db 257 KESRLDTQKRVDFLQMLDMSQSKETESHKALSDELVQAQSIIFIFAGYETTSVLSF 316
Qy 351 MEHRLSEDKAVQKLRREICQI--DTPMTLDDELNALPYLEAFVKESLRLLDPPSPYANRE 408
Db 317 IMVELATHPDVQKQLEIDAVLPNKAPPTDYVQLQMEYLDVMVNETLLEFPFIAMLERLV 376
Qy 409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRWL 468
Db 377 CKKQVE-----INGMFIPKGVVWVIMPSYALHRDP-KYWTEPEKFLPERFS 420
Qy 469 EDVTDLSNS-IEAPYGHQASFIISGPRACFGHFAVAEMKAEFLVTLRRVQREP 520
Db 421 KKNKDNIDPYIYTPFG-----SGPRNCIGMRFPALMMKMLALIRVLQNFSPRP 467

RESULT 5
US-09-957-997-3
; Sequence 3, Application US/09957997
; Patent No. US20020150915A1
; GENERAL INFORMATION:
; APPLICANT: Berkenstam, Anders
; APPLICANT: Bertilsson, Gran
; APPLICANT: Blomquist, Patrik
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-046001
; CURRENT APPLICATION NUMBER: US/09/957,997
; EARLIER FILING DATE: 2001-09-21
; EARLIER FILING DATE: 2000-09-22
; EARLIER FILING DATE: 2000-09-22
; EARLIER FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRI
; ORGANISM: Homo sapiens
US-09-957-997-3

Query Match
Best Local Similarity 29.0%; Score 385.5; DB 9; Length 503;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

Qy 13 LAAPSW--ASTAFSLYIAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLAVSLVLLLYGTHSHGLFKLGIPTLPFLGN---ALSFRKGYWTFDME 63
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLSNSTDPKVFNHVM-KEAYD---YKPGMAARVL 113
Db 64 CYKKYRKVWG-----IYDQQPMLAITDPMIKTVLKECYSVFTNRRPFGPVGFMKNAI 118
Qy 114 RIATGCGVWTAEGEAHRRHRRIMPSLSAQAVKSMVPIFLKGMELVDKMMEDAAEKDMA 173
Db 119 SI-----AEDEWKRIRLSLSPFTSGKLEKEMVPIIAQYGDVLRNLRREA----- 164
Qy 174 VGESAGEKATRLLET-EGVDVKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAVGLT 232
Db 165 -----ETGKPVTLKDVFGAYSDVITSTSGVNIIDSLNNPQD-----PFVENT 207
Qy 233 D---GPAPTLDSEKAIMWDFV-PYFRMK-RRHEIPTQGLAVS---RRVGIELMEQKQAV 286
Db 208 KLLRFNP-LDPPVLSIKVP-PFL-----TPILEALNITVFPKVISFTKSVQKIK 257
Qy 287 LGSASDAQVDKQVQGR--DILSLVRANIAPESQKLSDEEVLAQISNLLFAGYETS 344
Db 258 EGRL-----KETQKRVDFLQMLDMSQSKETESHKALSDELVQAQSIIFIFAGYETT 310
Qy 345 STVLTWVHRLSEDKAVQKLRREICQIDTM-----PTLDELNALPYLEAFVKESLRLLD 399
Db 311 SSVLSFTIYELATHPDVQKQVQKE---IDVLPENKAPPTDYVQLQMEYLDVMVNETLRLF 367
Qy 400 PPSPYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDA 459

```



Db 368 PVAMRLRVCKDVE-----INGMIPKGVVMIPSYVLHDP-KYWTPE 411  
QY 460 EBFPRPRLWEDVTDLSNS-IBAPYGHQASFSISGPRACFGWPAVAEMKAFVTLRLRVQF 518  
Db 412 EKFLPERFSKGNKNDIPYTPFG-----SGPRNCIGMRFALVNMKALVRVQLQNSF 465  
QY 519 EP 520  
Db 466 KP 467

RESULT 6  
US-10-425-114-38180  
; Sequence 38180, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 38180  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3608-017-E5\_FLI.pap  
US-10-425-114-38180

Query Match 13.5%; Score 384; DB 12; Length 537;  
Best Local Similarity 27.4%; Pred. No. 1.4e-25;  
Matches 162; Conservative 79; Mismatches 230; Indels 120; Gaps 24;  
QY 4 LVLTGALGLAFAFSWASIAFFS-LYAPRR-SSLNVLQGNHTNY--FTGNFLDILSART 59  
Db 23 LLIVYGVIG-ALLMKAALLDRLMWEPRRLRALRAOGLRGTYSRYLTG--DLREYR 78  
QY 60 GBHAKYRE-----KYGSTLRAGIAGAPVLNSTDPKVFNVMK 98  
Db 79 SKBEAWARPLRCHDIAGHVEFFIHGAVLEHGKTC-FTWFGVPRVTVPDLDARDMA 137  
QY 99 EAY-DYKPGMAARVIRIATGCGVVTAEAGEAHRHRRIMIPSLSAQAVKSMVPIFLEKGM 157  
Db 138 NKEGHEKPKFPA--LTKLFSDGVANHEGEKWKVXRRILNPAFHLEKLMPLAFSACCE 195  
QY 158 ELVDKMDAAEKDMAGVESAGEKATRLTEGVVDKDWGPRATLDVMALAGF-----D 211  
Db 196 ELVSRW-----AOSLGDGSCEL-----DVPDELQTLTGDIVISRTAFSGSYLEG 239  
QY 212 YKSDSLQKNTNELYAVFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPLTQGLAVS 271  
Db 240 RKIFQLOAEQAEMLSMII--DKFA-----VPGWMSLPTKNNRMRIKSEI 283  
QY 272 RRVGIELMEQKQAVLGSDAQVDKVDQGRDILSLVRANIANLPES-----QKLSDE 327  
Db 284 DSILRGLIGKRMQAMKQGESD-----KD-----DLGLLLESNARETGDSQSQPGGLTME 334  
QY 328 EVLAQISNLLFAGYETSSTVLTWMFHLRSEDKAVQDKLEEICQI--DTDMPTLDLNLAL 385  
Db 335 EVMECKLFFPAGNETTSVLTWTWVLLSWHPEDQDRAEEVLGFGKQKQGYDGLSRL 394  
QY 386 PYLEAFVKESRLRDPSPYANRECLKDEDPIPLAEPVIGRDSGVINEVRITKGTVMMLPLFNINRSKFIYGED 445  
Db 395 KTVTMILYELRLPYPAIAFSSRKYKE-----MVGVDVTPYAGVTLELPV 439

QY 446 FNINRSKFIYGEDAEERPERPRLWEDVTDLSNLSIEA--PYGHQASFSISGPRACFGWPAVA 503  
Db 440 LFIHDDPDWSDAHERPERPAEGVARASKDRALAFPPFGW-----GPRICIGQNFALL 493  
QY 504 EMKAPLFTVTLRRVQFEBIIISHPEYEHITLIISRRP--IVGREKEGYQMLQ 552  
Db 494 EAKWALSILORQFQF---LAPTPTVH-----PRRVIMLRPMHGAQIKLR 535

RESULT 7  
US-10-425-114-67055  
; Sequence 67055, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 67055  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4757-032-B12\_FLI.pap  
US-10-425-114-67055

Query Match 13.2%; Score 377.5; DB 12; Length 527;  
Best Local Similarity 24.9%; Pred. No. 5.3e-25;  
Matches 141; Conservative 106; Mismatches 212; Indels 107; Gaps 25;  
QY 6 LITGALGLAFAFSWASIAFFS---YLAPRRSSLNVLQGNHTNYFTGNFLDILSARTG-- 60  
Db 24 LLIGAL-----FFLVKPYVTVTRWFRGOGIGPSY-REFVGLSLEIKRMAAGS 71  
QY 61 -----EBHAKYREKYGSTURFAGIAGAPVLNSTDPKVFNVKWEAYD-Y 103  
Db 72 KITLDVGSHTDFVPIVQPYQYREWSDYKTFLY-WFGAVPTICVAEVLGVQVLAERTGLF 130  
QY 104 KPGMAARVIRIATGCGVVTAEAGEAHRHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKM 163  
Db 131 PKDYNDSS-MEVLGKGLVLANGEDWKRHEVHVHAPKPKDKLTMVVM-----ADLVQM 185  
QY 164 MEDAAEKDMAGVESAGEKATRLTEGVVDKDWGPRATLDVMALAGF--DYKSDSLQKNT 221  
Db 186 MQQWRSQ-----IQASNHEAE-IELSSSEFSELTSDVIAHTAFGTYSK-----EG 229  
QY 222 NELYAVFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPLTQGLAVSRVRGIELMEQ 281  
Db 230 KQVFA-----QKELQELTFTWIDIPAPAC-LKLPLTSKS---SRVV--EELDK 275  
QY 282 KQAVLGSDAQVDKQDVQ--GRDILSLVRANIANLPESQKLSDEEVLQISNLLFA 339  
Db 276 KVRSLMIAIEGRLAARGTSGVNDLLGLMLQAR-ALEQEGHQLMTTEIIVDECKTFIPA 334  
QY 340 GYETASTVLTWMFHLRSEDKAVQDKLEEICQIDTD-MPTLDLNLALPYLEAFVKESLRL 398  
Db 335 GQDTTSHLLTWTMFLLSRYSEWQRLREVURECGDAVNPDPVTVKULNMVLLSRL 394  
QY 399 DPSPYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMMLPLFNINRSKFIYGED 458  
Db 395 YSPVVFIR-----AVGSD-ILLSTRVFKGTMLSIPIALLHRDKDVWGQD 439  
QY 459 ABEFPERPRLWEDVTDLSL-----NSTEAPYGHQASFSISGPRACFGWPAVAEMKAFVTLR 514  
Db 440 ABEFNPDRFEHGSNAAAKHPNAL-----LSFSQGPRACTIGONFAMLEARIGIAMILQ 492

QY 515 RVQFEE--PIISHPEYEHITLIISRP 538  
Db 493 RFSELSNVVHAPKEAVTLM---PR 515

RESULT 8  
US-10-425-114-65616  
; Sequence 65616, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 65616  
; LENGTH: 547  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700100628\_FLI.pep  
US-10-425-114-65616

Query Match 13.2%; Score 376.5; DB 12; Length 547;  
Best Local Similarity 24.3%; Pred. No. 6.9e-25;  
Matches 144; Conservative 96; Mismatches 220; Indels 133; Gaps 25;  
QY 7 LTGALGLAASF--ASTAFSLYAPRRSSLYNLOQPNHTNY--FTGNELDILSARTGEE 62  
Db 37 LAGAVASVLLVLAWTLEMAWTPNLDRLRAQGLKGRYRLFTGLDRE--TAKVNR 94  
QY 63 -----HAKREKYGSTLRFAGIAGAPVLNSTDPKVFNVKAEYD 102  
Db 95 ARKPEPLGCHDITPRVQPMHSTIKYK-LSTFTWGTPTRVMI DPPELVKEVLSNKEG 153  
QY 103 Y---PKPQMAARVRIATGQDVVTABGAHKKRRHRIIMPISLSAQVKSVPFLEKGMEL 159  
Db 154 HFGKPRSSRIKGLL-----ANGLNVHDGKWKAKHRRILNPAFHHEKIKGMMPVSTCCIE 209  
QY 160 VDKWMDAEDKMDAVGESAGKATRLTEGVVDKMWVGRATLDVMAAGF--DYKSD-- 215  
Db 210 ITRW-----DMSMS-SEGSSB-----IDVWPEFQNLGDI SRTAFGSNYQEGRR 253  
QY 216 --SLQKNTNELYVAFVGLTDGFAPTLDSFKAIMMDFVP--YFTMKRRHEIPLTOGLAVS 271  
Db 254 IPFELQELAERLI-----QSVQTFIPGYWFLPTKNNRRM----- 288  
QY 272 RVVGLTEMEQKQAVLGSASDAQVDDKQVGR-----DILSLVRANI--AANLPESOKLS 325  
Db 289 RAIDVEIKRIURE-IIGKR-----EKDTKQRENTKDDLLGLLESNROSNGASLGLT 341  
QY 326 DEVLQAISNLLFAGYETSVLTWTFHRLSEDKAVQDKLREEI--CQIDTDMPTLDLNA 384  
Db 342 TEDVIEECKLFYFAGMETTSVLLTWTLVLSMHPWQERAREEVLSHFGRITTPDYDSL 401  
QY 385 LPYLEAFVKESLRDPPSPYANRECKDDEDFIPLAEPVIGRDSGVINEVRITKGTVMVLP 444  
Db 402 LKTVTMILYEVLRVLPVFLTRTYKEME-----LGGIKYPAAGVELLP 446  
QY 445 LFNINRSKFIYGEDAEFRPRERWLEDVTDLSNLSIEAPYGHQASF--GPRACFGWRFA 501  
Db 447 VFIHHDPDINGKDASENPEFP-----ANGISSATRHQAAPFPFGGPRICIGQSPA 499  
QY 502 VAEMKAFVTLVRRVQFE--PIISHPEYEHITLIISRPVIGREKEGYQMLQ 552

Db 500 LLEAKMTLCTILQRFSELSNVVHAPYVITL-----HPQGAQIRLK 543

RESULT 9  
US-10-425-114-61218  
; Sequence 61218, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 61218  
; LENGTH: 547  
; TYPE: PRT  
; ORGANISM: Oryza sativa nipponbare  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3431-054-B3\_FLI.pep  
US-10-425-114-61218

Query Match 12.9%; Score 368.5; DB 12; Length 547;  
Best Local Similarity 24.6%; Pred. No. 3.6e-24;  
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;  
QY 4 LVLLTGALGLAASFWSIASIAPFS-LYLAPRR-SSLVNLQGNHTNY--FTGNF----- 51  
Db 36 ILAAAAAVALLLAVSTLEWAWTPRRLRALRAQGIIRGNRYRLFTGDPVENVRLNR 95  
QY 52 -----LDLS-----ARTGEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVF 94  
Db 96 EARKPPLGCHDITPRVLPMSKAVEEGK-----PSFTWFGTPTRVWISDEIR 147  
QY 95 HVKKEA---YDYPKQMAARVRIATGQDVVTABGAHKKRRHRIIMPISLSAQVKSVP 151  
Db 148 EVMSNKGPHGYKPKTRLGKLL-----ASGVSVSEGEKWKAKHRRILNPAFHHEKIKRML 203  
QY 152 FLEKGMELVDKMDAEDKMDAVGESAGKATRLTEGVVDKMWVGRATLDVMAAGF- 210  
Db 204 FSNCCTEMVTRW-----ENSMST-SEGSSB-----VDVWPEFQNLGDI SRTAF 247  
QY 211 -DYKSD-----SLQKNTNELYVAFVGLTDGFAPTLDSFKAIMMDFVPYFTMKRRHEI 265  
Db 248 SSYEGRRIIFQLQAESAEIRI-----QAFRTI---FIPGYWFLPTKNNRRL 291  
QY 266 QGLAVSRVIGIELMEQKQAVLGSASDAQVDDKQVGRDILSLVRANI--AANLPESOKL 324  
Db 292 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLDGLLVESNRSNGKAELOM 342  
QY 325 SDEVLQAISNLLFAGYETSVLTWTFHRLSEDKAVQDKLREEIC-QIDTDMPTLDLNA 383  
Db 343 TTDIIIECKLFYFAGMETTSVLLTWTLVLSMHPWQERAREEVLSHFGRITTPDYDSL 402  
QY 384 ALPYLEAFVKESLRDPPSPYANRECKDDEDFIPLAEPVIGRDSGVINEVRITKGTVMWL 443  
Db 403 RLKVTVMILYEVLRVLPVFLTRTYKEME-----GGIKYPAEVT---LML 447  
QY 444 PLFNINRSKFIYGEDAEFRPRERWLEDVTDLSNLSIEAPYGHQASFIS---GPRACFGWR 500  
Db 448 PILFIHHDPDINGKDAEFPNPFADGICSNATK-----YQTSFFPFGGPRICIGQNF 500  
QY 501 AVAEMKAFVTLVRRVQFEPIISHPEYEH--ITLIISRPVIGREKEGYQMLQ 552  
Db 501 ALLEAKMAICTILQRFSELSNVVHAPYVITLHP-----OHGAQIKLK 545

```
RESULT 10
US-10-425-114-59349
; Publication No. US20040034888A1
; Sequence 59349, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59349
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-OSLELIB3474019G01_FLI.ppe
US-10-425-114-59349

Query Match      12.9%; Score 368.5; DB 12; Length 560;
Best Local Similarity 24.6%; Pred. No. 3.8e-24;
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLTGALGALAFSWASIAFFS-LYLAPRR-SSLYNLQGNHNY--FTGNF----- 51
Db 49 ILAAAAAVALLLAVSTLEWAWTPRRLERALLRAQGIKRNRYRLFTGDPVENVRLNR 108
QY 52 -----LDLS-----ARTGEHAKYREKYSTLRERAGIAGAPVLNSTDPKVEN 94
Db 109 EARKPLPLGCHDIIIPRVLPMSKAVEHGK-----PSTWFGPTPRVWISDPSIR 160
QY 95 HVKKEA---YDYPKPGMAARVLRIATGCVVTAEGEAHKKRRRIMPISLSAQAVKSMVPI 151
Db 161 EVMSNFGHYGKPKTRLGKLL---ASGVSYEGEKWAKHRRILNPAFHHEKIKRMLPV 216
QY 152 FLEKGMELVDKMDAEDAEKDMVAGESAGEKATRETEGVVDKVGKGRATLDVWALAGF- 210
Db 217 FSNCTEMVTRW-----ENSMISI-EGMSE-----VDVWPEFQNLTDGVISKTAFG 260
QY 211 -DYKSD----SLQKTNELNYVAFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPT 265
Db 261 SSYEGRIRIFQLQAESAEIRII-----QAFRTI---FIPGYWFLPTKNNRRLR 304
QY 266 QGLAVSRVGIELMEQKQAVLGSASDQAVDKVDKQGRDILSLVRANI-AANLPESOKL 324
Db 305 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLLGLLVESNMRESNGKAE LGM 355
QY 325 SDEVLQAIISNLLFAGYETSSVLTWTFHRLSDEKAVQDKLREEIC-QIDTDMPTLDEN 383
Db 356 TTDBIIECKLFYFAGMETTSVLLTWLIVLSMHPWQERAREEVLLHFGRTTPDYDSLS 415
QY 384 ALPYLEAFVKESLRDPPSPYANRECKQBDFTPLAEPVIGRDSGVINEVRITKTMWML 443
Db 416 RLKIVTMILYEVLRLYPPVFLTRTYKEMEL-----GGIKYPAEVT-----LML 460
QY 444 PLFNINRSKFIYGEDAEERPERWLEVDVDSLSIEAPYGHQASFIIS---GPRACFGWRF 500
Db 461 PILFIHDPDINGKADGAFNPGFADGISNATK-----YQTSFPFGWGPRICIGQNF 513
QY 501 AVAEMKAFLVTLRVRVQFEPIISHPEYEH--ITLIISRPVIGREKGYQMRLQ 552
Db 514 ALLEAKQAICTILQRFSE---LSPSYIHAPFTVITLHP-----QHGAQIKLK 558
```

```
RESULT 11
US-10-425-114-59350
; Sequence 59350, Application US/10425114
```

```
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59350
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSFLCYP143F07_FLI.ppe
US-10-425-114-59350
```

```
Query Match      12.9%; Score 368.5; DB 12; Length 562;
Best Local Similarity 24.6%; Pred. No. 3.8e-24;
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLTGALGALAFSWASIAFFS-LYLAPRR-SSLYNLQGNHNY--FTGNF----- 51
Db 51 ILAAAAAVALLLAVSTLEWAWTPRRLERALLRAQGIKRNRYRLFTGDPVENVRLNR 110
QY 52 -----LDLS-----ARTGEHAKYREKYSTLRERAGIAGAPVLNSTDPKVEN 94
Db 111 EARKPLPLGCHDIIIPRVLPMSKAVEHGK-----PSTWFGPTPRVWISDPSIR 162
QY 95 HVKKEA---YDYPKPGMAARVLRIATGCVVTAEGEAHKKRRRIMPISLSAQAVKSMVPI 151
Db 163 EVMSNFGHYGKPKTRLGKLL---ASGVSYEGEKWAKHRRILNPAFHHEKIKRMLPV 218
QY 152 FLEKGMELVDKMDAEDAEKDMVAGESAGEKATRETEGVVDKVGKGRATLDVWALAGF- 210
Db 219 FSNCTEMVTRW-----ENSMISI-EGMSE-----VDVWPEFQNLTDGVISKTAFG 262
QY 211 -DYKSD----SLQKTNELNYVAFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPT 265
Db 263 SSYEGRIRIFQLQAESAEIRII-----QAFRTI---FIPGYWFLPTKNNRRLR 306
QY 266 QGLAVSRVGIELMEQKQAVLGSASDQAVDKVDKQGRDILSLVRANI-AANLPESOKL 324
Db 307 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLLGLLVESNMRESNGKAE LGM 357
QY 325 SDEVLQAIISNLLFAGYETSSVLTWTFHRLSDEKAVQDKLREEIC-QIDTDMPTLDEN 383
Db 358 TTDBIIECKLFYFAGMETTSVLLTWLIVLSMHPWQERAREEVLLHFGRTTPDYDSLS 417
QY 384 ALPYLEAFVKESLRDPPSPYANRECKQBDFTPLAEPVIGRDSGVINEVRITKTMWML 443
Db 418 RLKIVTMILYEVLRLYPPVFLTRTYKEMEL-----GGIKYPAEVT-----LML 462
QY 444 PLFNINRSKFIYGEDAEERPERWLEVDVDSLSIEAPYGHQASFIIS---GPRACFGWRF 500
Db 463 PILFIHDPDINGKADGAFNPGFADGISNATK-----YQTSFPFGWGPRICIGQNF 515
QY 501 AVAEMKAFLVTLRVRVQFEPIISHPEYEH--ITLIISRPVIGREKGYQMRLQ 552
Db 516 ALLEAKQAICTILQRFSE---LSPSYIHAPFTVITLHP-----QHGAQIKLK 560
```

```
RESULT 12
US-09-992-901-2
; Sequence 2, Application US/09992901
; Patent No. US20020073446A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M.
```

APPLICANT: Chory, Joanne  
TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING  
TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING  
FILE REFERENCE: SALKINS.024DV1  
CURRENT APPLICATION NUMBER: US/09/992,901  
CURRENT FILING DATE: 2001-11-14  
PRIORITY APPLICATION NUMBER: US 09/527,073  
PRIORITY FILING DATE: 2000-03-16  
PRIORITY APPLICATION NUMBER: US 60/124570  
PRIORITY FILING DATE: 1999-03-16  
PRIORITY APPLICATION NUMBER: US 60/170,931  
PRIORITY FILING DATE: 1999-12-14  
PRIORITY APPLICATION NUMBER: US 60/172,832  
PRIORITY FILING DATE: 1999-12-20  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-992-901-2

Query Match 12.9%; Score 367; DB 9; Length 520;  
Best Local Similarity 22.2%; Pred. No. 4.6e-24;  
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;  
QY 4 LVLLTGLGIAAFSASIAFFSLYLAPR-----SSLYNLQGNHNTYFTGNFLDLS--- 56  
DB 13 VLVLSTLVSVKMSL-----LWPRKIEHFHFKQIRGPPY-HFFIGNVKELVGMML 67  
QY 57 -----ARTGEHAKRYEKYSTLRFAAGIAGAPVNST--DPKVFNNHMKEA 100  
DB 68 KASSHPMPFPHSHNLPRVLSFYHWRKIYATFL--VWPGPFRLTADVADPOLIREIFS 124  
QY 101 YDPKPGMAARVRIATGCVTTAECAHRRRIMISLSAQAVKSMVPIFEKGMELV 160  
DB 125 EFYEK-NEAHLVKQLEGGLSLKGEKWAHRKIIISPTFHENKLLVPPVLSVTDNV 183  
QY 161 DKWMDAAEKDVAVGSAGEKATRLTEGVDYKQVGVGATLDVMAAGFDYKSDSLQNK 220  
DB 184 DKMSDKLSEN-----GEVE-----VDVYENFQILTEDVISRTAFGSSVE----- 222  
QY 221 TNELYAVFGLTGFAP-----TLDSKAINMDVPPVFRMTKRHEIPTQGLAV 270  
DB 223 -----DGRAVFLQAQMLLCAEAFQV---FIPGYRFP-----TRGNLK 260  
QY 271 SRRVG-----TELMEQKQAVLGSASQAVDKQVQGRDILSLVRANIAANLPESQK 323  
DB 261 SRKLDKEIRSLKLLERRQNAIDGEGEC---KEPAKDLLGLMIOA-----KN 308  
QY 324 LSDEEVLQAINLLFAGYETSTVLTWFMHRLSEDKAVQDKLREICQI--DTMPTLDE 381  
DB 309 VTVDIVIEECKSPFFAGKQTTSLTLLTSMHPPEWQAKARDEVIRVCGSRDVPYTKDH 368  
QY 382 LNALPVLEAFVKSRLRDPSPVANRECLKDEDFIPLABPVRGDSVINEVRIKTMV 441  
DB 369 VVKUKTLSMINESLRYPPIVATIRAKSDVK-----LGGYKIFCGTEL 413  
QY 442 MLPLFNIRSKFIYGEDAEFRPERMLEVDYDLSNIEAPYGHQASFTS---GPRACFGW 498  
DB 414 LIPIIAVHDDQAIKNGDWNENFARFADGVPRAAK-----HPVGFIFFGLGVRTCIGQ 466  
QY 499 RFAVAEMKAFVTLRVRQVEPIISHPEYEHITLII 534  
DB 467 NLAIQAKLTUAVMIQRTFH---LAPTYOHAPTVL 499

RESULT 13  
US-10-425-114-38827  
Sequence 38827, Application US/10425114  
Publication No. US2004003488A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 38827  
LENGTH: 662  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700045236\_FLI.pep  
US-10-425-114-38827  
Query Match 12.8%; Score 365.5; DB 12; Length 662;  
Best Local Similarity 22.9%; Pred. No. 9.1e-24;  
Matches 123; Conservative 96; Mismatches 193; Indels 125; Gaps 16;  
QY 70 YGSTRFAGIAGAPVNSTDPKVFNNHMKAEYDPKPGMAARVRIATGCVVTAEGEAH 129  
DB 188 YGGIFRL-NFGPKSLIYSDPAIAKHILRENSKAYSKGILAEILFEVMTGLIPADGEIW 246  
QY 130 KHRHRIMIPSLSAQAVKSMVPIFEKGMELVDKMDAAEKDVAVGSAGEKATRLTE 189  
DB 247 RYRRRAIYPAHLHOKVYTAIIGLGEASORLCEKL-----DKAA--VDGE 288  
QY 190 GVDVMDVGRATLDVMAAGFDYKSDSLQNTNELYAVFVGLTDPGAPTLDSPKAINWDF 249  
DB 289 DMEMESLFSRLTLDVIGKAVFNYPDLSYDNGIVEAVVYTLRE--AEMRSTSPITW-- 344  
QY 250 VPYFTMKRHEIPTQGLAVSRVGIELMEQKQAVLGSASQD-----AVDKQDVQ 302  
DB 345 -----EIPWKDISPRQKVNALK-----LINSTLDELIAICKELVEQEDLQF 388  
QY 303 RD-----ILSLLVRANIAANLPESQKLSDEVLQAINLLFAGYETSTVLTWFMH 353  
DB 389 HBEYMNEQPSILHFL-----LASGDDVSSKQLRDDLMTMLIAGHETSAAVLTWTFY 440  
QY 354 RLSEDKAVQDKLREICQIDTD-MPTLDLALPVLEAFVKSRLRDPSPVANRECLKD 412  
DB 441 LSKYKPVNAKQDRAADVGLGDLTIEDVKLKYTRVINESRLYPOPPVLIERSLED 500  
QY 413 EDFIPLAEFVIGRDSVINEVRIKTMVMLPLFNIRSKFIYGEDAEFRPERMLEVDVT 472  
DB 501 D-----ILGGYPIGRGEDIFISVNLHCP-KHWDAAEVFNPERWPLDGP 544  
QY 473 DLSNIEAPYGHQASFTSGPRACFGWRFAVEMKAFVTLRVRQVE----- 519  
DB 545 NP-NEINQNFYSY-LPFGGGRKCVGDMFATFETVVAATMLVYKRFDFQMAPGAPPVDMTG 602  
QY 520 -----PIISHPEYEHIT-----LIISRPYVIREKEG 546  
DB 603 ATHTTEGLKMTVTRTPPVIPNLEMKIITDSQBSTLSAPSMVVSASVASGEDQG 659  
RESULT 14  
US-10-114-270-60  
Sequence 60, Application US/10114270  
Publication No. US20040030110A1  
GENERAL INFORMATION:  
APPLICANT: Guo, Xiaojia  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Miller, Charles E.  
APPLICANT: Malyankar, Uriel M.  
APPLICANT: Spyttek, Kimberly A.  
APPLICANT: Patturajan, Meera  
APPLICANT: Liu, Ziaohong  
APPLICANT: Gusev, Vladimir Y.

```
APPLICANT: Li, Li
APPLICANT: Vernet, Corine
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Gorman, Linda
APPLICANT: Shenoy, Suresh G.
APPLICANT: Pena, Carol E.A.
APPLICANT: Smithson, Glennda
APPLICANT: Burgess, Catherine E.
APPLICANT: Gerlach, Valerie
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shinkets, Richard A.
APPLICANT: Gangolli, Esha A.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Casman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Liette, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shlomit R.
APPLICANT: Stone, David J.
APPLICANT: MacDougall, John R.
APPLICANT: Rothenberg, Mark E.
APPLICANT: TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
APPLICANT: FILE REFERENCE: 21402-322C
APPLICANT: CURRENT APPLICATION NUMBER: US/10/114,270
APPLICANT: CURRENT FILING DATE: 2002-11-27
APPLICANT: PRIOR FILING DATE: 2001-04-03
APPLICANT: PRIOR APPLICATION NUMBER: 60/281,086
APPLICANT: PRIOR FILING DATE: 2001-04-03
APPLICANT: PRIOR APPLICATION NUMBER: 60/281,136
APPLICANT: PRIOR FILING DATE: 2001-04-03
APPLICANT: PRIOR APPLICATION NUMBER: 60/281,863
APPLICANT: PRIOR FILING DATE: 2001-04-05
APPLICANT: PRIOR APPLICATION NUMBER: 60/281,906
APPLICANT: PRIOR FILING DATE: 2001-04-05
APPLICANT: PRIOR APPLICATION NUMBER: 60/282,020
APPLICANT: PRIOR FILING DATE: 2001-04-06
APPLICANT: PRIOR APPLICATION NUMBER: 60/282,930
APPLICANT: PRIOR FILING DATE: 2001-04-10
APPLICANT: PRIOR APPLICATION NUMBER: 60/282,934
APPLICANT: PRIOR FILING DATE: 2001-04-10
APPLICANT: PRIOR APPLICATION NUMBER: 60/283,512
APPLICANT: PRIOR FILING DATE: 2001-04-12
APPLICANT: PRIOR APPLICATION NUMBER: 60/283,710
APPLICANT: PRIOR FILING DATE: 2001-04-13
APPLICANT: PRIOR APPLICATION NUMBER: 60/284,234
APPLICANT: PRIOR FILING DATE: 2001-04-17
APPLICANT: Remaining Prior Application data removed - See File Wrapper or PALM.
APPLICANT: NUMBER OF SEQ ID NOS: 470
APPLICANT: SEQ ID NO 60
APPLICANT: LENGTH: 520
APPLICANT: TYPE: PRT
APPLICANT: ORGANISM: Homo sapiens
US-10-114-270-60

Query Match 12.8%; Score 365; DB 12; Length 520;
Best Local Similarity 24.8%; Pred. No. 6.9e-24;
Matches 143; Conservative 104; Mismatches 221; Indels 114; Gaps 23;

Qy 2 FILVLLTGALGLAA--FSWASIAFFSLYLAPRRSLYQLQ---GNHNTNYTGNFLDILS 56
Db 19 WLLLLLVGASWLLARILAMT---YTFY-----DNCRLQCPFPKQWFWGHLGLVTP 69
Qy 57 ARTGEH-----AKYREKY-----GSTLRGAGTAGAPVLNSTDPKVFNVHVMKEAYDPKPG 107
Db 70 TEEGMKTLTQVLYTQGGFKWGLPIPIVICHPTDIRSI-----TNSAAIVPKDN 122
Qy 108 MAARVLRIATGCGVTTAEGBAHKRHRIRIMPSLSAQAVKSMVPIFLEKGMELVDKQMEDA 167
Db 123 LFYSFLKPLWGLDGLLLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKQHLA 182
Qy 168 AEKDMVAGESAGEKATRLTEGVVDKVDWVGRATLDVMAAGFDYKSDSLQNKTNELVYA 227
Db 183 SE-----GSSR-----LDMPHFISLMTLDSLOKCVFSPES-NCEKPS-E-YIA 223
```

```
Qy 228 FVGLTDFAPLTDSFKAIMWDFVYF-----RTMKRRHEIPLTQGLAVSRR-----V 274
Db 224 AIELSAFVEKRNQOILLHIDFLYLTPTDQGRPHRACRLVHD--FTDAVIQERRTLPSQ 281
Qy 275 GIELMEQKQAVLGSASDAQVDKDKVQGRDILSILVRANIANNLPSSQKLSDBEEVLAQIS 334
Db 282 GVDDFLQAK-----AKSKTLDFIDV-----LLSKD-----EDGKLSDEDIRAED 323
Qy 335 NLLFAGYETSTVLTWMPHRLSEDKAVQDKLREBEICQIDTDMPTL-----DELNALPYLEA 390
Db 324 TFMPEGHDTTASGLSVLYHLAKHPEYQECRCQEVQELLKDRPEFIEMDDLAQLPFLTM 383
Qy 391 FKESLRLDPPSPYANRECKLDEDFIPLAEPVIGRQGSVINEVRITKGTWMLPLENINR 450
Db 384 CIKESLRHLPVPVVISRGCTQD-----FVLPGRV-----IPKGLICLSVGTTH 429
Qy 451 SKFIYGEDAEFERPERWLEDVDTSLNSIEAPYGHQASFISSGRACFGWRFPAVAMKAFIF 510
Db 430 NPAVM-PDEVYDPFRF-----DPENIKERSPLAFIPFGAGPRNCIGQTFAMAEMKVLA 483
Qy 511 VTLRVQPEPIISHPEVEHITLIISRPRIYGREKEGVQWMLQ 552
Db 484 LTLRFRVLPDHTPR-----RKPELVIRAEGLWLRVE 517

RESULT 15
US-10-424-599-199559
; Sequence 199559, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199559
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22226C.i.pap
US-10-424-599-199559

Query Match 12.4%; Score 353; DB 12; Length 524;
Best Local Similarity 24.6%; Pred. No. 8.5e-23;
Matches 145; Conservative 89; Mismatches 215; Indels 140; Gaps 25;

Qy 1 MFILVLLTGALGLAAFPWSIAFFSLYLAPRRSL-----YNLOGPNHNTNYTGNFLDILS 56
Db 13 VFVIVI-----LALTSWAMRLNMLWRPKRLERLLREQGLQG-NPYRILVGDGLKEIVK 65
Qy 57 ARTGEHAK-----YREKYGSTLR-----FAGTAGAPVLNSTDPKVFNVHVMKE 99
Db 66 LQM-EARSKPNLNSHDIVPRFAHLHOSVLKHGKNSIFWFGPKPRVTLTDPGLIKDVLNK 124
Qy 100 AYDYPKPGVAARVLRIATGCGVTTAEGBAHKRHRIRIMPSLSAQAVKSMVPIFLEKGMEL 159
Db 125 ISDFPKPEANPLAKILAT--GLVNYDGEKWNKRRRLINPAFSLKILMILPIFFKSCNDL 182
Qy 160 VDKWVEDAAEKDMVAGESAGEKATRLTEGVVDKVDWVGRATLDVMAAGFDYKSD----- 215
Db 183 IIRK-----EGMUSYDSCSE-----MDVWPFQNLASDVIAARTAFGSSFEGKR 226
Qy 216 --SLQNKTNELVYAFVGLTGDFAPTLDSFKAIM-----WDFVPYFPTMKRHEIPLTQ 267
Db 227 IFQLQKELAEI-----TMKVMKVYIPGWRFVP-TATNRRMKEI----- 264
```

Qy	268	LAVSRVGIEMEQKQAOVLGSASQAVDKDVQGRDILSLVRA-----NIA	315
Db	265	---DRYIKASLTDMIKK-----REKAPKTGEATRDDLLGILLESNHKEIOEHRNNENVG	315
Qy	316	ANLPESQKLSDEVLQAQISNLLFAGYETSTVLTWMFHLSEDKAVQDKLREEICQI-DT	374
Db	316	MNI-----NDVIECKLFYFAGQETTSVLLVWTVLLSRYPDWQSRAREEVLQVFGK	367
Qy	375	DMPTLDELNALPYLE-AFVKESLRLDPPSPYANRECKDEDEFIPLAEPVIGRDSVINEV	433
Db	368	QAPNFDGLSHLKIWTMIFLKVILRYPPAVGLNRN-----VDRDMKLG-N-L	412
Qy	434	RITKGTWMLPLFNINRSKFIYGEDAEERPERWLEDVTDLSN:IEA--PYGHQASFISSG	491
Db	413	SLPAGVQVSLPTTMVPHDRELWGDVNEFKPERFSEGVLKATNGRVSPFPFCM-----G	466
Qy	492	PRACFGWRFAVAENKAFLFVTLRRVQFEPFIISHPEYEH--ITLIISRPR	538
Db	467	PRICIGQNFSLLEAKMALSTILOHFSPE---LSPAYAHAPVTFTLQPO	512

Search completed: April 2, 2004, 14:10:37  
Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17 Seconds

(without alignments)  
3151.687 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852

Sequence: 1 MFILVLLTGALGLAFAFWAS.....RIVGREKEGYQMLQKPV 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78: \*\*  
1: P11:  
2: P12:  
3: P13:  
4: P14:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	419	14.7	593	2 F86441	probable cytochrom
2	406	14.2	503	2 JC4702	cytochrome P450 3A
3	404	14.2	503	2 S14275	steroid 6beta-mono
4	402	14.1	501	2 A34235	cytochrome P450 3A
5	400	14.0	526	2 T02191	cytochrome P450 ho
6	399	14.0	504	2 A29410	cytochrome P450, S
7	398.5	14.0	503	2 A29815	cytochrome P450 3A
8	398	14.0	503	2 S50211	cytochrome P450 3A
9	395.5	13.9	503	2 S28168	nifedipine oxidase
10	394	13.8	504	2 A60564	cytochrome P450 3A
11	385.5	13.5	503	2 JX0062	cytochrome P450 3A
12	385	13.5	502	1 A34101	cytochrome P450 3A
13	379	13.3	501	2 A29487	cytochrome P450 3A
14	370	13.0	504	2 A22631	cytochrome P450 3A
15	368	12.9	504	2 S08092	cytochrome P450 3A
16	365	12.8	520	2 H84663	probable cytochrom
17	363.5	12.7	511	2 T00864	cytochrome P450 ho
18	362.5	12.7	523	2 B96662	probable cytochrom
19	359.5	12.6	511	2 S66472	cytochrome P450 4B
20	357	12.5	502	2 JX0334	cytochrome P450 3A
21	355.5	12.5	512	2 A96695	hypothetical prote
22	350.5	12.3	1054	1 A49975	NADPH-ferrihemopro
23	350	12.3	526	2 JC4533	cytochrome P450 4F
24	349.5	12.3	511	1 O4HUB1	cytochrome P450 4B
25	349.5	12.3	524	2 S29723	cytochrome P450 4F
26	346.5	12.1	511	1 B40164	cytochrome P450 4B
27	346.5	12.1	524	2 JC7594	cytochrome P450 en
28	346.5	12.1	524	2 JC7598	cytochrome P450 en
29	346	12.1	537	2 T02450	probable cytochrom

30	345.5	12.1	520	2 T24778	hypothetical prote
31	345	12.1	516	2 T00514	cytochrome P450 ho
32	344	12.1	520	1 A46661	leukotriene B4 ome
33	343	12.0	504	2 A25222	cytochrome P450 3A
34	340.5	11.9	524	2 T09944	probable cytochrom
35	340	11.9	517	2 T02192	probable cytochrom
36	339.5	11.9	508	2 T16980	probable cytochrom
37	338.5	11.9	503	2 JC7627	cytochrome P450 3A
38	337.5	11.8	524	2 T09999	cytochrome P450 -
39	337	11.8	520	2 S45702	leukotriene-B4 20-
40	336.5	11.8	490	2 T00404	probable cytochrom
41	336	11.8	522	2 JC4532	hypothetical prote
42	334.5	11.7	509	2 T24785	cytochrome P450 4F
43	334.5	11.7	576	1 H71414	probable cytochrom
44	333.5	11.7	497	2 S57097	cytochrome P450III
45	333.5	11.7	499	2 T18699	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

F86441

probable cytochrome P450 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 20-Apr-2001

C/Accession: F86441

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: F86441

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-593 <STO>

A/Cross-references: GB:AE005172; NID:g11136728; PIDN:AAG31309.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology

C/Keywords: heme; iron; metalloprotein

F,514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match

Best Local Similarity 14.7%; Score 419; DB 2; Length 593;

Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NFDILSATG--EHHAKYREKVGST-----LRFAGI-----AGAPVLN 86

DB 93 NVLDFFWDTGSDQYPKVPEAKGSIQAVRNEAFPIYELFLTYGGIFLITGPKSFLI 152

QY 87 STDPKVFNVHMK--AYDPKPGMAARVRIATGCVTVABCEAHKRRHRRIMISLSAQV 145

DB 153 VSDPSIAKHILKDNKAYSK-GILAEILDVFMGKGLIPADGEIWRRRRAIVPALHOKYV 211

QY 146 KSMVPFFLEKGMELVDKMMEDAAEKMAVGESAGEKKATLETGVVDKDWVGRATLDVM 205

DB 212 AMISLFGASDELQCKL--DAA-----ALKGEVEMESLPSRLTLDII 253

QY 206 ALAGFYKSDSLQNTNELYVAVGLTDPGAPTLDSPKAIMDVPVFRFWKRRHEIPLT 265

DB 254 GKAVFYDPSLNTDNTGVIEAVYTVLREADRVSPIP--VWD-IPIWKDIS-----PRQ 305

QY 266 QGLAVRRRGIELMEOKKQAVLGASDAQVKKDVQ-----GRD--ILSLVRANIAA 316

DB 306 RKVATSLKINTLDD-----LIATCKMWEESLQPHFEYMWNERDPSILHFL----- 353

Query March	14.2%	Score 406;	DB 2;	Length 503;
Best Local Similarity	27.2%;	Pred. No. 6.5e-19;		
Matches 144;	Conservative 86;	Mismatches 179;	Indels 120;	Gaps 22;
54	ILSARTG-EZHAKY-REKYGSTLRFAGI--AGAPVLNSTDPKFNHYM-KEAY-----	101		
50	ILAYKGFWEFDYCHHKYG--KLWGLYDGRQPLAITDPDIITKVLVKECYSTFTNRR	106		
102	DYPRKGMAARVLRIATGDGVWTAEGSAHKKRRRIMIPLSLAQAVKSVVPILEKGMELVD	161		
107	NFGPVGLLKKAISI-----SEDEWKIRPALLSPFTTSGLKENEFPII-----	149		
162	KWEDAAAEKDMAVGESAGEKKAIRLETEGVYKDWVGRATLDWMALAGFDYKSDSLONKT	221		
150	NQYTDMLVRNMROGSEEGP-----TSMKDIFGAYSMOIVATISFGVNDVSLNNPQ	200		
222	NELYVAFVGLTDGFATLDSFKAIM-WD-FVYPFTMKRRHEI-PLTQGLAVSR-----	272		
201	D-----PFVEKVKLLKFDIFDFLFLSVTLFPFLPLFEALNVSMFPFRDVI	246		
273	--RVGTELMEQKKQAVLGASDAQVDDKDVCGGRDILSLVVRANIAANLPESOKLSDEEV	329		
247	DFFKTSVERMKENR-----MKEKEKQRMDFLQLMINSQNSKVKDSHKAISDVEI	295		
330	LAQISNLLFAGYETSSTVLTMFHRUSEKXAVQDKLREEICQIDTDP-----TLBELNA	384		
296	VAQSVITFAGYETTSALSFVLYLAIHPDIQKQLQDE---IDAALFNKAHAFTYDILLQ	352		
385	LPYLEAFVKSLRLDPPSPYANRECLKDBDFIPLAEPVIGRDGGSVINEVRIKGTWYMLP	444		
353	MEYLDWMVNETLRLPYTAGLERVCKTDE-----INGVFTPGTVMVMP	397		
445	LFNINRSGFYGEDABEFPERVLEVDVDTSLNSIAPYGHQASFTISGPRACFGWRPAVAE	504		





```

A:Title: Isolation and sequence determination of a cDNA clone related to human cytochrome P-450
A:Reference number: A25517; MUID:87041402; PMID:3464943
A:Accession: A25517
A:Molecule type: mRNA
A:Residues: 1-391,'W','K',393-503 <BBA>
A:R:Book, R.W., Muto, T., Besune, P.H.; Srivastava, P.K.; Lloyd, R.S.; Guengerich, F.P.
J. Biol. Chem. 264, 910-919, 1989
A:Title: Characterization of mRNA species related to human liver cytochrome P-450 nifed
A:Reference number: A32199; MUID:89093163; PMID:2463251
A:Accession: A32199
A:Molecule type: protein
A:Residues: 'X',2-9,'XXX',13-15 <BOR>
A:R:Komori, M.; Hashizume, T.; Ohi, H.; Miura, T.; Kitada, M.; Nagashima, K.; Kamataki, T.
J. Biochem. 104, 912-916, 1988
A:Title: Cytochrome P-450 in human liver microsomes: high-performance liquid chromatogr;
A:Reference number: PX0012; MUID:89214010; PMID:3243766
A:Accession: PX0012
A:Molecule type: Protein
A:Residues: 'X',2-11,'X',13-25 <XOM>
A:Experimental source: Liver microsome
C:Genetics:
A:Gene: GDB:CYP3A4
A:Cross-references: GDB:118782
A:Map position: 7q22.1-7q22.1
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metall;
F:302-464/Domain: Cytochrome P450 homology <P45>
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 398.5; DB 2; Length 503;
Best local Similarity 28.7%; Pred. No. 2e-18;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAPFSW--ASIAFSLYLAPRES--SLYNLQG-PNHNTNY-FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLAVSLVLLYLXGTHSHGLFKXGLGPGPTLPPLGN--ILSYHKGFQCFMDE 63

QY 62 EHAKYREKYGSTLRFAGIAGAVLNSTDPKVFNVHW-KEADYD---PKPGMAARVLRIAT 117
DB 64 CHKKYKGVKWF---YDQ---QQPVLAITDPDMIKTVLVKECVSYFTNRRPPGPGVTFMKSA- 117

QY 118 GDGVVTAEGSAHKRRHRIIMPISLQAQVKSVPVIFLEKGMELVDQWMDAAEKDQMAVGES 177
DB 118 ---ISIAEDEMKRLSLSPTFTSGKLKEMVPIIAOYGDVLVNLAREA----- 164

```

178 AGEKATRIET-EGVDVKDVGKATLDVMAAGFYKSDSLQNTKTNELIYAVFGLTDGPA 236  
165 -----ETGKPVTLKDFVGYSDMDVITSTSFSGNIDSLNNPD----- 201  
237 PTLDSFKAM-WDFV-PYERTMK-RRHEIPLTQGLAV---SRRVGIELMEOKQOAVLGS 290  
202 PFVENTKLLRPDPLDFPLSTVFPFLPILEVLNLCVFPREVTFNLKRSVKR-----M 256  
291 SDQAVKDVQGRDILSLIVRIANIANLPESOKLSDBEVLQAI SNLLPAGVETSTVLTW 350  
257 KESRLDTQGRVDFLQMLDSDNSKSTESHKALSDELVQAQSIIIFIPAGVETSTVLSF 316  
351 MFHRLSDKAVDKLREEICQI--DMDPTLDELNALPYLEAFVKSRLRLDPPPPYANRE 408  
317 IMTELATHPDVQKQLQEDDAVLNPKAPETYDTVLMQMEYLDVMVNETLRLPFIAMLERV 376  
409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMPLFNINRSKFIYGEDAEERPERWL 468  
377 CKXDVE-----INGMFIPKGVVVMIPSYALHRDP-KYWEPEKFLPERFS 420  
469 EDVTDLSNS-IEAPYGHQASFISSPRACFGWRFAVAKAEFLVTLRRVQPEP 520  
421 KKKNDIDPIYTPFG-----SGPRNCIOMRFPALMMKMLIRVLQNFESFKP 467  
RESULT 8  
S50211  
cytochrome P450 3A13 - mouse  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Mus musculus (house mouse)  
C>Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-Jul-2000  
C/Accession: S50211; S18155  
R:Yanagimoto, T.; Itoh, S.; Sawada, M.; Hashimoto, H.; Kamataki, T.  
Biochim. Biophys. Acta 1201, 405-410, 1994  
A>Title: Molecular cloning and functional expression of a mouse cytochrome P-450 (Cyp3a-  
A/Reference number: S50211; MUID:95101705; PMID:7803471  
A/Accession: S50211  
A/Molecule type: mRNA  
A/Residues: 1-503 <YAN>  
A/Cross-references: EMBL:X63023; NID:G50634; PIDN:CAA44754.1; PID:G50635  
C/Genetics:  
A/Gene: CYP3A13  
C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
F:302-464/Domain: cytochrome P450 homology <P45>  
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted  
Query Match 14.0%; Score 398; DB 2; Length 503;  
Best Local Similarity 27.0%; Pred. No. 2.2e-18;  
Matches 142; Conservative 74; Mismatches 179; Indels 130; Gaps 20;  
QY 63 HAKYREKYGSTLRPAGIAGAPVLNSTDPKVNHYM-KEAYD-----YPKPGMAARVLR 114  
DB 65 HKYKGMWG---LYDG--RQPVLAITDPDIKTVLVKECYSTFTNRRRFGVGLKKALS 119  
QY 115 IATGQGVVTAEGAHRHRRIMIPSLAQAVKSMVPIFLEKGMELVDKMMEDAAEDMAV 174  
DB 120 I-----SENEEKRIRALLSTFTSGRLKEMFPII-----NQFTDVLVNRMQ 162  
QY 175 GESAGEKATRIETEGVDVKDVGKATLDVMAAGFYKSDSLQNT----- 221  
DB 163 G-LGEGKPT-----SMKIDIFGAYSDMTATSFSGNIDSLNNPDQPFVEKIKLLKF 213  
QY 222 ---NELYAVFGLTDGFAPTLDSFKAIW--DFVPYERTMKRHEIPLTQGLAVSRVGI 276  
DB 214 DIFDPLFLS-VTLFPFLTFVFDALNLSLFPDRDVISFTT-----SVERKEN 259  
QY 277 ELMEQKQAVLGSASDAQVDKQVGRDILSLVRIANIANLPESOKLSDBEVLQAI SNL 336  
DB 260 RMKEKEKQV-----DFLQMINSONYKTKESHKALSDEVEIQAQVIF 302  
QY 337 LFAGVETSTVLTWFMHRLSEDKAVODKLRBEICQIDTDMF-----TLDLNLALPYLEAF 391

303 IPAGVETTSALSPALYLLAIHPDVKQLQDE---IDAALPNKAPATYDTLLQMEYLDV 359  
QY 392 VKESIRLPPPPYANRECLKXDEDEFIPLAEPVIGRDSGVINEVRITKGTVMPLFNINRS 451  
DB 360 VNETHRLPYIAGRIERLVCKTVE-----INGLFIPKGTVMIPFALHKD 404  
QY 452 KFIYGEDAEERPERLEWEDVTDLSNSIEAPYGHQASFISSPRACFGWRFAVAKAEFLV 511  
DB 405 P-KYWEPEERPERFSKKNQDSIN---PYMY-LPFGSGPRNCIGMRFALINMKVALVR 459  
QY 512 TLRRVQFPI-----ISHREYEHITLIISRPRIVGRE 543  
DB 459 VLQNTVQPCKEITPLKSKQGLLOPENPLLLKVSARDETVDSE 503  
RESULT 9  
S28168  
nifedipine oxidase (EC 1.14.14.-) cytochrome P450 3A8 - crab-eating macaque  
N/Alternate names: cytochrome P450 CMLc; cytochrome P450 MKnf2; cytochrome P450-MK2  
C/Species: Macaca fascicularis (crab-eating macaque)  
C/Date: 28-May-1993 #sequence\_revision 28-May-1993 #text\_change 28-Jul-2000  
C/Accession: S28168; S94509; S36875  
R:Komori, M.; Kikuchi, O.; Sakuma, T.; Funaki, J.; Kitada, M.; Kamataki, T.  
Biochim. Biophys. Acta 1171, 141-146, 1992  
A>Title: Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity of the pri  
A/Reference number: S28168; MUID:93129612; PMID:1282830  
A/Accession: S28168  
A/Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-503 <KOM>  
A/Cross-references: GB:S53047; NID:G263692; PIDN:AAB24952.1; PID:G263693  
R:Ohta, K.; Kitada, M.; Hashizume, T.; Komori, M.; Ohi, H.; Kamataki, T.  
Biochim. Biophys. Acta 996, 142-145, 1989  
A>Title: Purification of cytochrome P-450 from polychlorinated biphenyl-treated crab-eat  
A/Reference number: S04509; MUID:89287352; PMID:2500151  
A/Accession: S04509  
A/Molecule type: protein  
A/Residues: 1-22 <OHT>  
R:Omori, S.; Horie, T.; Guengerich, F.P.; Kiuchi, M.; Kitada, M.  
Arch. Biochem. Biophys. 305, 405-413, 1993  
A>Title: Purification and characterization of two forms of hepatic microsomal cytochrome  
A/Reference number: S36874; MUID:93364294; PMID:8373178  
A/Accession: S36875  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-13 <OHM>  
C/Genetics:  
A/Gene: CYP3A8  
C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
F:302-464/Domain: cytochrome P450 homology <P45>  
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted  
Query Match 13.9%; Score 395.5; DB 2; Length 503;  
Best Local Similarity 28.6%; Pred. No. 3.2e-18;  
Matches 155; Conservative 85; Mismatches 187; Indels 115; Gaps 26;  
QY 13 LAAFSW--ASIAFSSLYLAPRRS-SLYNIQG-PNHNY-FTGNFLDILSARTG-----E 61  
DB 7 LAVETWLLAVTLVLLVLYGTHSHGLFKKLGIPGPTPLPLGN---ILSYRKGFWTFDME 63  
QY 62 BHAKYREKYGSTLRPAGIAGAPVLNSTDPKVNHYM-KEAYD-----YPKPGMAARVLR 113  
DB 64 CYKYGKVMGF---YDG--RQPVLAITDPNMTKVLVKECYSVFTNRRRFGVFGMKNAI 118  
QY 114 RIATGQGVVTAEGAHRHRRIMIPSLAQAVKSMVPIFLEKGMELVDKMMEDAAEKQVA 173  
DB 119 SI-----AEDEWKIRISLLSPTSGKLKEMVPIIAKYGDVLRNLNREA----- 164  
QY 174 VGESAGEKATRIET-EGVDVKDVGKATLDVMAAGFYKSDSLQNTKTNELIYAVFGLT 232  
DB 165 -----ETGKPVTLKDFVGYSDMDVITSTSFSGNIDSLNNPD----- 201



C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metalloprotein; cytochrome P450 homology <P45>  
F:302-464/Domain: cytochrome P450 homology <P45>  
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.5%; Score 385.5; DB 2; Length 503;  
Best Local Similarity 29.0%; Pred. No. 1.4e-17;  
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

QY 13 LAAFSW--ASIAFFSLYL-APRRSLYNLQ-PTHNY-FTGNFLDILSARTG-----E 61  
DB 7 LAVETWLLAVSLILLYLGYTRTHGLFKLIGPGPTPLFLGN---ALSFRKGYWTFDE 63

QY 62 EHAKEYREKYSTLRPAGIAGAPVLNSTDPKVFNVHM-KEAYD-----YKPGMAARVL 113  
DB 64 CYKRYKRWG-----YDCCQPMALITDPMIKTVLVKECYVFNTRPPFPVGMKMAI 118

QY 114 RIATGDGVWTAEGEAHRRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDMA 173  
DB 119 SI-----AEDBEWKIRSLSPFTSGKLKEMVPIIAQYGDVLVRLRREA----- 164

QY 174 VGSAGEKKATRLT-EGVDVQKDWGRATLDVVALAGFDYKSDSLQNTNELYVAFVGLT 232  
DB 165 -----ETGKPVTLKHVFGAYSMVDVITSTSGVSDSLNNPQD-----PFVENT 207

QY 233 D---GPAFTLDSFKAIMWDFVFRTRMKRHEIPLTQGLAVS---RRVGIELMEQKQAV 286  
DB 208 KLLRFPN-LDPVLISKVF-PFL-----TPILEALNIVFRKVISFLTKSVKQIK 257

QY 287 LGSASDAQVKKQVGR--DILSLVRANITANLPSQKLSDEEVLQAISNLLPAGYETS 344  
DB 258 EGRL-----KETQHRVDVFLQMLIDSONSKSDSETHKALSLELMAQSIIFIFAGYETT 310

QY 345 STVLTWFMHRLSEDKAVQDKLREICQIDTM-----PTLDELNALPYLEAFVKESLRLD 399  
DB 311 SSVLSFIIYELATHPDVQKQVKE---IDVLFNKAPPYDVTVLQVLELDVWVNETLRF 367

QY 400 PPSPYANRECLKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFYIGSDA 459  
DB 368 PVAMRLERVCKQVE-----INGMFIPKGVVWVMPISVYLHDP-KYWTETP 411

QY 460 EEPFPERKLEVDTSLS-IEAPYGHQASFISSGPRACFGWRFAVAEMKAFVTLRRVQF 518  
DB 412 EKFLPERFSKKNIDPIYITPFG-----SGPRNCIGRMFALVNMKALVRVLQNTSF 465

QY 519 EP 520  
DB 466 KP 467

RESULT 12  
A:34101  
cytochrome P450 3A5 - human  
N:Alternate names: cytochrome P450 HLP2  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: A34101; S06491; I52302  
R:Aoyama, T.; Yamano, S.; Waxman, D.J.; Lapenson, D.P.; Meyer, U.A.; Fischer, V.; Tyndal  
J. Biol. Chem. 264, 10388-10395, 1989  
A:Title: Cytochrome P-450 hPCN3, a novel cytochrome P-450 IIA gene product that is diff  
DNA-expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and cyclosporine.  
A:Reference number: A34101; MUID:89278095; PMID:2732228  
A:Accession: A34101  
A:Molecule type: mRNA  
A:Residues: 1-502 <AOY>  
A:Cross-references: GB:J04813; NID:g181345; PIDN:AAA02993.1; PID:g181346  
R:Schuetz, J.D.; Molowa, D.T.; Guzelian, P.S.  
Arch. Biochem. Biophys. 274, 355-365, 1989  
A:Title: Characterization of a cDNA encoding a new member of the glucocorticoid-responsi  
A:Reference number: S06491; MUID:90025114; PMID:2802615  
A:Accession: S06491  
A>Status: not compared with conceptual translation

A:Molecule type: mRNA  
A:Residues: 1-304; P:306-317, F:319-323, D:325-376, G:378-502 <SCH>  
R:Jounaidi, Y.; Guzelian, P.S.; Maurel, P.; Villarejo, M.J.  
Biochem. Biophys. Res. Commun. 205, 1741-1747, 1994  
A:Title: Sequence of the 5'-flanking region of CYP3A5: comparative analysis with CYP3A4  
A:Reference number: I52302; MUID:95110318; PMID:7811260  
A:Accession: I52302  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-24 <RES>  
A:Cross-references: GB:S74699; NID:g786472; PIDN:AA14157.1; PID:g4261857  
C:Gene: CYP3A5  
A:Cross-references: GDB:118783  
A:Map position: 7q22.1-7q22.1  
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monoxygenase;  
F:302-463/Domain: cytochrome P450 homology <P45>  
F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.5%; Score 385; DB 1; Length 502;  
Best Local Similarity 28.1%; Pred. No. 1.5e-17;  
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAFSW--ASIAFFSLYL-APRRSLYNLQ-PTHNY-FTGNFLDILSARTG-----E 61  
DB 7 LAVETWLLAVSLILLYLGYTRTHGLFKLIGPGPTPLFLGN---VLSYRQGLWKPDTE 63

QY 62 EHAKEYREKYSTLRPAGIAGAPVLNSTDPKVFNVHM-KEAYDYPKGMARVLRATG-- 118  
DB 64 CYKRYKRWG---YEG--QLPVLATDPDVRTVLVKECY-----SVFTNRSRLGPV 111

QY 119 ---DGVWTAEGEAHRRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKMAV 174  
DB 112 GFMKSAISLAEDBEWKIRSLSPFTSGKLKEMVPIIAQYGDVLVRLRRE-AEKG--- 167

QY 175 GESAGEKKATRLTETGVQKDWGRATLDVVALAGFDYKSDSLQ-----KTNELYVAF 228  
DB 168 -----KPVTLKIDFGAYSMVDITGTSFGVNIIDSLNNPQDPFVESIKKFLK 213

QY 229 VGLTGDFAPTLDSFKAIMWDFVFRTRMKRHEIPLTQGLAVS---RRVGIELMEQKQAV 286  
DB 214 ---GFLDPL-FLSIIPLPFL-----TFVFEALNVSLFPKDTINFLSKS----- 252

QY 287 LGSASDAQVKKQVGRDILSLVLRANITANLPSQKLSDEEVLQAISNLLPAGYETSST 346  
DB 253 VNRKKSRLNDKQKHLDFLQMLIDSONSKSETESHKALSLELAQSIIFIFAGYETSS 312

QY 347 VLTWFMHRLSEDKAVQDKLREICQI---DTPMTLDELNALPYLEAFVKESLRDPPSPY 404  
DB 313 VLSFTLVELATHPDVQKQKLEIDAVLPNKAPPTVDVAVQMEYLDVWVNETLRLFFVAIR 372

QY 405 ANRECLKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFYIGEDAEEFRP 464  
DB 373 LERTCKQVE-----INGVFIPKGSVVIPTVYALHDP-KYWTETEEFRP 416

QY 465 ERWLEDVTDLSLS-IEAPYGHQASFISSGPRACFGWRFAVAEMKAFVTLRRVQEP 520  
DB 417 ERFSSK-KDSIDPIYITPFG-----TGFPRNCIGRMFALVNMKALIRVLQNFSEKP 466

RESULT 13  
A:429487  
cytochrome P450 3A6 (version 1) - rabbit  
N:Alternate names: cytochrome P450 3c  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text\_change 28-Jul-2000  
R:Dalet, C.; Clair, P.; Daujat, M.; Fort, P.; Blanchard, J.M.; Maurel, P.  
DNA 7, 39-46, 1988  
A:Title: Complete sequence of cytochrome P450 3c cDNA and presence of two mRNA species w  
A:Reference number: A29487; MUID:88166352; PMID:3349903

J. Biochem. 107, 718-725, 1990

A;Title: Purification and characterization of four catalytically active testosterone 5 $\alpha$ -nally related forms.

A;Reference number: PX0032; MUID:90375438; PMID:2398038

A;Accession: PX0035

A;Molecule type: protein

A;Residues: 1-26 <NAG>

A;Experimental source: liver, Sprague-Dawley male rat

R;Lechner, M.C.

submitted to the EMBL Data Library, December 1991

A;Reference number: S21697

A;Accession: S21697

A;Molecule type: mRNA

A;Residues: 1-206 'A', 208-212, 'I', 214-231, 'V', 233-504 <IEC>

A;Cross-references: EMBL:X64401; NID:G56038; PIDN:CRA45743.1; PID:G56039

R;Sibairo, V.; Lechner, M.C.

Arch. Biochem. Biophys. 293, 147-152, 1992

A;Title: Cloning and characterization of a novel CYP3A1 allelic variant: Analysis of CYP3A1

A;Reference number: S36137; MUID:92117688; PMID:1731631

A;Accession: S36137

A;Molecule type: mRNA

A;Residues: 205-206 'A', 208-212, 'I', 214-231, 'V', 233-234 <RIB>

A;Cross-references: EMBL:X64401

R;Telhada, M.B.; Pereira, T.M.; Lechner, M.C.

Arch. Biochem. Biophys. 298, 715-725, 1992

A;Title: Effect of dexamethasone and phenobarbital on run-on transcription rate and CYP3A1

A;Reference number: S27107; MUID:93037516; PMID:1417000

A;Accession: S27107

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-24 <TEL>

A;Cross-references: EMBL:X62085

R;Cooper, K.O.; Reik, L.M.; Jayyosi, Z.; Bandiera, S.; Kelley, M.; Ryan, D.E.; Daniel, J.

Arch. Biochem. Biophys. 301, 345-354, 1993

A;Title: Regulation of two members of the steroid-inducible cytochrome P450 subfamily (CYP3A1 and CYP3A2)

A;Reference number: S30378; MUID:93213168; PMID:7681660

A;Accession: S30378

A;Molecule type: protein

A;Residues: 1-25 <COO>

R;Burger, H.

Proc. Natl. Acad. Sci. U.S.A. 90, 2145-2149, 1992

A;Title: Paradoxical transcriptional activation of rat liver cytochrome P-450 3A1 by de to primary monolayer cultures of adult rat hepatocytes.

A;Reference number: I59218; MUID:92196074; PMID:1372436

A;Accession: I59218

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-24 <BUR>

A;Cross-references: GB:M86850; NID:G205919; PIDN:AAA41780.1; PID:G205920

C;Genetic:

A;Gene: CYP3A1, P450P

C;Superfamily: human cytochrome P450 CYP3A5, cytochrome P450 homology

C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase

F;303-465/Domain: cytochrome P450 homology <P45>

F;443/Binding site: heme iron (Cys) (axial ligand) #status Predicted

Query Match 13.0%; Score 370; DB 2; Length 504;

Best Local Similarity 26.2%; Pred. No. 1, 5e-16;

Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY 7 LTGALGUAAPSWSIASFFSLY-IAPRRSLYNLQG-----PNTNTFTGNP-LDI 54

Db 3 LLSALTLETWLLAVLVLLYLGFTGTHGLFKQGIQPKPLPFGTVLNYMGLWKFV 62

QY 55 LSATGTGEHAKYREKYSTLRFAGIAGAVLNSTDPKVFNEWM-KEAY-----DYPKP 106

Db 63 -----ECHKYKING-----LFDG--QMLFAITTEMINKVLYKECSVFTNERDFGPV 111

QY 107 GMAARVLRATGQVTVTAGEAHKRRIRNITPSLAQAQKSNVIFLEKGMELVDKMD 166

Db 112 GI-----NGKAVSVAKDEEKRYALLSPFTSGRLKEMFPIIQYGDILVYLKQE 163

QY 167 AAEKMDVAGSAGEKKATRLT-EGVDVQDWGVRATLDVMALAGFDYKSSLOW----- 219

```
Db 164 A-----ETGKPVTKKVGAGYSDVITSTSPGVNVDLSLNNPKDPV 204
QY 220 -KTNEL-----YVAFVGLTDGFAPTLDSFKAINW--DFVPYPRTWKGRHEIPL 264
Db 205 EKTKLLRFDFFDLFVSVLFPFLT-----PIVEMLNICMFPKDSIEFFK----- 250
QY 265 TQGLAVSRVIELMEQKQAVLGASDAQVKKDVQGRDILSLVRANIAANLPESQ-K 323
Db 251 -----KFYRKETRLDSVOKHRV-----DFLQIMNAHNSDKKESHTA 290
QY 324 LSDEVLQAISNLLFAGYETSSVLTWMFRLSDEKAVQDKLREBICQI--DTDMPTLDE 381
Db 291 LSDMEITAQSIIFIFAGYETSSVLTWMFRLSDEKAVQDKLREBICQI--DTDMPTLDE 381
QY 382 LNALPYLEAFVKESLRLDPPSPYANRCLDKDEDPIPLAEPVIGRDGSGVINEVITKGTWV 441
Db 351 VMEYELDMVNLNETLRLPIGNRLRVCCKDVE-----INGVMPKGSVV 395
QY 442 MLPLFNINRSKFIYGEDAEPRPERWLEDVTDLSNS-IEAPYGHQASFGSPRACFGWRF 500
Db 396 MIPSVLALHRDQHWPE-PEERPRPERFSKENKGSIDPVVLPFG-----NGPRNCIGMRF 448
QY 501 AVAENKAFVTLRRVQPEP 520
Db 449 ALMNMKALATKVLQNFSPQ 468

RESULT 15
S0892
cytochrome P450 3A16 - mouse
N:Contains: oxidoreductase (SC 1.1.1.1)
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 09-May-1997 #text_change 28-Jul-2000
C:Accession: S0892
R:Itch, S.; Satch, M.; Abe, Y.; Hashimoto, H.; Yanagimoto, T.; Kamataki, T.
Eur. J. Biochem. 226, 877-882, 1994
A:Title: A novel form of mouse cytochrome P450 3A (Cyp3a-16). Its cDNA cloning and expression
A:Reference number: S0892; MUID:95112853; PMID:7813478
A:Accession: S0892
A:Molecule type: mRNA
A:Residues: 1-504 <ITO>
A:Cross-references: EMBL:D26137; NID:g493670; PIDN:BAA05133.1; PID:g666968
A:Experimental source: fetal liver
A>Note: In the authors' translation residues 9-20 do not match the nucleotide sequence
A>Note: the authors translated the codon TTC for residue 464 as Glu
C:Genetics:
C:Gene: CYP3a-16
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
F:303-465/Domain: cytochrome P450 homology <P45>
F:469/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 12.9%; Score 368; DB 2; Length 504;
Best Local Similarity 25.8%; Pred. NO. 2e-16;
Matches 132; Conservative 100; Mismatches 171; Indels 108; Gaps 21;

QY 68 EXYGTSL-REFAGIAGPVLNSTDPVFNHVM-KEAY-----DYPKPGMAARVLRATG 118
Db 66 EXYGTWGLFDG--QIPLEVTIDPETIKNLVKECFSVFTNRQDFPPVGMKSIISL--- 120
QY 119 DGVTVAEGEBAHRHRIRIIPSLAQAVKSMVPIFLKGMELVDMKMDAAEKQAVGESA 178
Db 121 -----AKDEWKRYRALLSPFTTSGNLKEMFPVIEQYG-DILVKYLRQEAEGKQFVA--- 171
QY 179 GEKATRLTEGVVDKDWVGRATLDMALAGFDYKDSLQNKTNELYVAVGLTDGFAPT 238
Db 172 -----VKDVLGAYMDVITSTTFGVNIDSLNPNP-----PF 203
QY 239 LDSFKAIM-WDF-----VPYFRTWKRRHEIPLAQGLAVSRVGIELMEQKQAVLGSA 290
Db 204 VENAKVLAFLDFDPLSLVALFPFLTPYIEM--LNICMFPKDSIEFFK---FVDRM 256
```

```
QY 291 SDOAVDKDVQGRDILSLVRANIAANLPESQK-LSDEVLQAISNLLFAGYETSSVLT 349
Db 257 TENLDSKQKHRVDFFYLAMEAVNKSQKDSHKALSEIETITQSIIFIFAGYETSSILS 316
QY 350 WMFRLSDEKAVQDKLREBICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANR 407
Db 317 FTVYSLATHPDIOKQLOEBIDEALPNKAPPTYDTVMAMEYELDMVNLNETLRLPIINRLO 376
QY 408 ECLKDEDPIPLAEPVIGRDGSGVINEVITKGTWVMLPLFNINRSKFIYGEDAEPRPERW 467
Db 377 VCKDVE-----INGYIPKGSVTVIIPSVVLHHDQHWPE-PEEFQPERF 420
QY 468 LEDVTDLSNS-IEAPYGHQASFGSPRACFGWRFVAENKAFVTLRRVQPEPIISHPE 526
Db 421 SKENKGSIDPVVLPFG-----NGPRNCIGMRFALMNMKALIKVLQNFSPQPC----K 470
QY 527 YEHTTLIISPRIVGREKEGYQMRLQ-VKPV 556
Db 471 ETQIPKLSR-----ELLQPVKPI 490
```

Search completed: April 2, 2004, 14:01:28  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 11 Seconds  
(without alignments)  
2636.643 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852

Sequence: 1 MFILVLTGALGAFAFWAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414	14.5	503	1 CP3T_PIG	P79401 sus scrofa
2	413	14.5	503	1 CP3B_RAT	P51869 rattus norv
3	408.5	14.3	503	1 CP3O_SHEEP	Q29496 ovis aries
4	404	14.2	503	1 CP3C_CAMEL	P24463 canis fam
5	402	14.1	501	1 CP36_RABIT	P11707 oryctolagus
6	399	14.0	503	1 CP33_HUMAN	P05184 mus musculu
7	398	14.0	503	1 CP3D_MOUSE	Q64464 mus musculu
8	396.5	13.9	502	1 CP34_HUMAN	P08684 mus musculu
9	395.5	13.9	503	1 CP38_MACFA	P33268 macaca fasc
10	394	13.8	504	1 CP3B_MOUSE	Q64459 mus musculu
11	391	13.7	504	1 CP31_MOUSE	Q91ma7 mus musculu
12	385.5	13.5	503	1 CP37_HUMAN	P24462 homo sapien
13	385	13.5	502	1 CP35_HUMAN	P20815 homo sapien
14	385	13.5	507	1 CP3S_BOVIN	P79102 bos taurus
15	377.5	13.2	503	1 CP3L_CALJA	O18993 callithrix
16	371.5	13.0	501	1 CP3V_MESAU	O70537 mesocricetu
17	370	13.0	504	1 CP31_RAT	P04800 rattus norv
18	368	12.9	504	1 CP3G_MOUSE	Q64481 mus musculu
19	362	12.7	503	1 CP3F_CAVPO	Q84466 cavia porce
20	359.5	12.6	511	1 CP4B_MOUSE	Q84462 mus musculu
21	356.5	12.5	496	1 CP30_FUNHE	Q9pve8 fundulus he
22	355.5	12.5	503	1 CP3E_CAVPO	Q64417 cavia porce
23	351.5	12.3	496	1 CP36_FUNHE	Q8axv5 fundulus he
24	351	12.3	520	1 CP58_HUMAN	P98187 homo sapien
25	350.5	12.3	1054	1 CP5E_BACSU	O08336 bacillus su
26	350	12.3	504	1 CP32_RAT	P05183 rattus norv
27	350	12.3	526	1 CP35_RAT	P51870 rattus norv
28	349.5	12.3	524	1 CP31_RAT	P33274 rattus norv
29	348.5	12.2	511	1 CP4B_HUMAN	P13584 homo sapien
30	346.5	12.1	511	1 CP4B_RAT	P15129 rattus norv
31	346.5	12.1	524	1 CP3C_HUMAN	Q9hcs2 homo sapien
32	345.5	12.1	520	1 YR2_CAMEL	Q27514 caenorhabdi
33	344.5	12.1	502	1 C340_ORYLA	Q98t91 oryzias lat

34	344	12.1	520	1 CP33_HUMAN	Q08477 homo sapien
35	344	12.1	524	1 CP3B_HUMAN	Q9bbi6 homo sapien
36	340.5	11.9	524	1 CP72_CATRO	Q05047 catharacthu
37	340	11.9	503	1 CP3P_MOUSE	Q09158 mus musculu
38	338.5	11.9	503	1 C343_HUMAN	Q9hb55 homo sapien
39	337	11.8	496	1 C4AB_DROME	O46054 drosophila
40	337	11.8	503	1 CP3H_CAVPO	Q64409 cavia porce
41	337	11.8	520	1 CP32_HUMAN	P78329 homo sapien
42	336.5	11.8	518	1 CP3R_ONCMY	O42563 oncorhynchu
43	336	11.8	522	1 CP34_RAT	P51869 rattus norv
44	334.5	11.7	509	1 YRVA_CAMEL	Q27516 caenorhabdi
45	334.5	11.7	580	1 C973_ARATH	O23365 arabidopsis

#### ALIGNMENTS

##### RESULT 1

ID	CP3T_PIG	STANDARD;	PRT;	503 AA.
AC	P79401;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome P450 3A29 (EC 1.14.14.1) (CYPIIA29).			
GN	CYP3A29.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Small intestine;			
RX	MEDLINE=98347363; PubMed=9682441;			
RA	Nissen P.H., Winteroe A.K., Fredholm M.;			
RT	"Mapping of porcine genes belonging to two different cytochrome P450 subfamilies";			
RL	Anim. Genet. 29:7-11 (1998).			
CC	!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.			
CC	!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.			
CC	!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.			
CC	!- INDUCTION: P450 can be induced to high levels in liver and other tissues by various foreign compounds, including drugs, pesticides, and carcinogens.			
CC	!- SIMILARITY: Belongs to the cytochrome P450 family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; Z93099; CAB07513.1; ..			
DR	HSSP; F14779; IUPZ.			
DR	InterPro; IPR001128; Cytochrome_P450.			
DR	InterPro; IPR008072; EP450_CYP3A.			
DR	Pfam; PF00067; P450; 1.			
DR	PRINTS; PR01689; EP450IICYP3A.			
DR	PRINTS; PR00385; P450.			
DR	PROSITE; PS00086; CYTOCHROME_P450; 1.			
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;			
KW	Microsome; Endoplasmic reticulum.			
FT	METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).			
SQ	SEQUENCE 503 AA; 57198 MW; E3D411B2674FD17F CRC64;			
Query Match	14.5%;	Score 414;	DB 1;	Length 503;



```

Best Local Similarity 27.4%; Pred. No. 8.9e-17;
Matches 158; Conservative 85; Mismatches 197; Indels 136; Gaps 23;

QY 12 GLAASWASIA--FSLYLAPERS-----SLNLOQPNHTNYFTGNELDILSARTGEH-- 63
Db 6 GSTETWULATLSVLVLYGYTHGLFKKIGIPGRLPYF-GN---ILGRKGVDFD 61
QY 64 AKYREKYGSTLRFAGI--AGAPVLNSTDPKVNHYM--KEAY-----DYPKPGMAARVL 113
Db 62 KXCFOQYG--KMGVYDGRQLLAATVDPNMLKSLVKECYSVFTNRRSFGPLGAMENAL 118
QY 114 RTATGCVVTAAGEAHKRRIMWLSAQAVKSMVPIFLEKGMELVDKXMEDAARDMA 173
Db 119 SL-----ADDEWKRIITLSLFTTSGKLEKEMFPIISHYGDLLVSNLAKE-AEG-- 167
QY 174 VGESAGEKKATLETGVDGVKDWGRATLDVNALAGPDYKSDSLQNK-----TNELVAF 228
Db 168 -----KPVTKMDIFGAYSMVDVITSTAFGVNIDSLNNPDPPFVNSKKLLK 212
QY 229 VGLTDGFAPTLDSFKALMDF-----VPYF-RTMKRHEIPTQGLAVSRVG 275
Db 213 FSFPDFPLSLIFFPPLTFPIFVNLITLFPKSSVNFFTKSVKMKESLT----- 262
QY 276 IELMEOKQOAVLGASDAQVKKDQVGRDILSLVRANIAANLPSQKLSDEVLQISN 335
Db 263 -----DQQRV-----DLLQLMINSONKEMDPKLSNEELVAQGI 301
QY 336 LIFAGVETSTVLTWPHRLSDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAFVK 393
Db 302 FIFAGVETSSALLAYELATHDPVQKLOBEIATFNKAPPYDALAQMEVLDVWN 361
QY 394 ESRLDPPSPYANRECLDKEDFIPLAEPVIGRDGVSINEVRIKGTMTWMLPLFNINRSKF 453
Db 362 ETLLPLPIAARLACKDVE-----IHGVFPKGTVVVVFVLRDPD 406
QY 454 IYGEDAEERPERLWEDVDLSNIEAPYCHQASISGPRACFGWRFAVEMKAFVTL 513
Db 407 LWPE-BEEFPERPFRSKKHDTIN---PYTY-LPFGTGRNCGMRFALMNKLALVRVL 460
QY 514 RYQVFEPI-----ISHPEYEHITLIISR 536
Db 461 QNFSFKCKETQIPLKLTQGLTQPKPVVLKILPR 496

RESULT 2
CP39 RAT
ID CP39 RAT STANDARD; PRT; 503 AA.
AC P51538; Q64557; Q64631.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome P450 3A9 (EC 1.14.14.1) (CYP11A9) (P450-OLF3) (Olfactive)
DE (3AH15).
GN CYP3A9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97144501; PubMed=8990268;
RA Mainke A., Strockamp D., Roos P.H., Hanstein W.G., Chabot G.G.,
RA Nef P.;
RT "Expression and inducibility of cytochrome P450 3A9 (CYP3A9) and
RT other members of the CYP3A subfamily in rat liver.";
RL Arch. Biochem. Biophys. 337:62-68 (1997).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96220175; PubMed=8660329;
RA Wang H., Kawashima H., Strobel H.W.;
RA "cDNA cloning of a novel CYP3A from rat brain.";
RT

```

```

Biochem. Biophys. Res. Commun. 221:157-162 (1996).
CC FUNCTION: This isozyme seems to be implicated in olfaction. Active
CC in the demethylation of erythromycin as well as benzphetamine.
CC CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC TISSUE SPECIFICITY: Mainly expressed in olfactory epithelium.
CC SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U60085; AAB03662.1; -;
CC EMBL; U46118; AAC52582.1; -;
CC PIR; JC4702; JC4702.
CC HSP; P14779; 1JPZ.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008072; E450_CYP3A.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PRO1689; EP450ICYP3A.
CC PROSITE; PS00385; P450.
CC PROSITE; PS0086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC Microsome; Endoplasmic reticulum; Olfaction.
CC METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC CONFLICT 442 457 F -> V (IN REF. 2).
CC SEQUENCE 503 AA; 57811 MW; 9BBB13E690675EB4 CRC64;
CC
Query Match 14.5%; Score 413; DB 1; Length 503;
Best Local Similarity 27.4%; Pred. No. 1e-16;
Matches 145; Conservative 86; Mismatches 178; Indels 120; Gaps 22;

QY 54 ILSATG-EEHAKY-REXGSLTFAGI--AGAPVLNSTDPKVNHYM--KEAY----- 101
Db 50 ILAYRKGFWFEDKYCHKYK--KLWGLYDGRQVLAITDPDIKTVLVEKCSFTNRR 106
QY 102 DYPKPGMAARVLRIATGDGVVTAGEAHKRRHRIIMPSLSAQAVKSMVPIFLEKGMELVD 161
Db 107 NFGPVGILKALSI-----SEDEWKRIALLSPTFTSGKLKEMFPII----- 149
QY 162 KWMDAAEKDQAVGESAGEKKATLETGVDVQVQVGRATLDVNALAGFYKSDSLQNK 221
Db 150 NQYTDMLVRNMRQSGSEKGP-----TSMKDFGAYSMVDVITATSGVNVDSLNNPQ 200
QY 222 NELYVAFVGLTDGFAPTLDSFKALM-WD-FVPYFRTMKRRHEI-PLTQGLAVSR----- 272
Db 201 D-----PFVEKVKLLKFDIFDPLFLSVTLFPFLTFLPEALNVSMFPRDVI 246
QY 273 ---RVGIELMEOKQOAVLGASDAQVKKDQVGRDILSLVRANIAANLPSQKLSDEEV 329
Db 247 DPFKTSVERMKENR-----MKEKQRQMDFLQLMINSONSKYKDSHKAISLVEI 295
QY 330 LAQISNLLPAGYETSTVLTWPHRLSDKAVQDKLREICQIDTDMF-----TLDELNA 384
Db 296 VAQSVIFIPAGYETTSALLSVFLVLLAIHPDIQKLODE---IDAAPNKAHATYDILLQ 352
QY 385 LPYLEAFVKESLRLDPPSPYANRECLDKEDFIPLAEPVIGRDGVSINEVRIKGTMTWMLP 444
Db 353 MEYLDVMVNETLRLYPIAGRLERLVCKTDE-----INGVFIPKGTVMVIP 397
QY 445 LFNINRSKFIYGEDAEERPERLWEDVDLSNIEAPYCHQASISGPRACFGWRFAV 504
Db 398 TFAHKDHP-YWPEEERPERPFRSKKHQNNIN-----PTMY-LPFGNGPRNCGMRFALMN 451
QY 505 MKAFLEVTLLRRVQFEPI-----ISHPEYEHITLIISRPIV 540
Db 452 MKVALFRVLQNTSFQCKETQIPLKSLQGLLOPKLPKLLKLVSRDET 500

```

```

RESULT 3
CP30 SHEEP STANDARD; PRT; 503 AA.
ID CP30 SHEEP STANDARD; PRT; 503 AA.
AC Q29496;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A24 (EC 1.14.14.1) (CYP3A24)
GN CYP3A24.
OS Ovis aries (Sheep).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Ching M.S., Chun-Jing J., Ghabrial H., Wooley P.J., Smallwood R.A.,
RA Morgan D.J.;
RT "Ovine foetal liver CYP3A24."
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U59378; AB02657.1; -
CC HSP; P14779; IJFZ.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008072; EP450_CYP3A.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01689; EP450ICYP3A.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57360 MW; 804223EAD0304238 CRC64;
Query Match 14.3%; Score 408.5; DB 1; Length 503;
Best Local Similarity 28.5%; Pred No. 1.8e-16;
Matches 165; Conservative 82; Mismatches 197; Indels 135; Gaps 27;
7 LTGALGLAFAFWASIAFFSLYLAPRRS-----SLYNLQGNHNYFTGNFLDILSARTG-- 60
3 LIPSPSLETWLLALSLVLLVLYGYSHGLFKLGVSQPRLPYF-GN---VLSYRGVC 58
61 -----EHAHYREKYSTLRERAGIAPVLNSTDPKFNHYM-KEAYD-----YKPGM 108
59 EFDEECFKYKGQWGV-----FEG--KQPLVITDPTVITLVKRCYGVFNRRVFGPMGI 113
109 AARVLRIATGCVVTAEGEAKHRRIRIMPSLSAQAKVMVPIFEKGMELVDMMEDAA 168
114 MK-----NAVSVAEDEQWKIRITLLSPFTSGKLKDMFPIIGKYGDVLVRLAKE-A 164
169 EKDMVAGSAGEKKATRLTEGTVDVQKDWVGRATLDVWALAGFDYKSLQNKINELYVAF 228

```

```

Db 165 EK-----GKS-----VNMKDFGAVSMDVITSTSGVNIIDSLGNP----- 199
Qy 229 VGLTDFGAPT-----LDSF--KATWDF-VPYER---TMKRHEIP-LTQGLAVS 271
Db 200 ---QDFVENAKLLRNFLDPFLSVLFLPFIPEVLNITMPPKSAVDFLIK--SVK 254
Qy 272 RRVGIELMEQKQAVLGSASDAQVDKQVQGRDILSLVRANIANLPESOKLDEEVL 331
Db 255 RIKESRLKDNQPRV-----DFQLMINSONSKETDNHKLALSQELMA 297
Qy 332 QISNLLFAGYTSSTVLTWFRHLSKXAVQKLREEICQIDTDM-----ETLDELNALP 386
Db 298 QSVIFIFAGYETTSNTLSLLYLATHPDVQOKLOE---IDATPPNKAPTYDVLAOME 354
Qy 387 YLEAFVKESLRDPPSPYANRECKDEDFIPLAEPVIGRDSVINEVRITGTVMVLELF 446
Db 355 YLDMVNVETIRVFPIAIVRLDLCKKDE-----IHGVSIPKGTAVTVEIF 399
Qy 447 NINRSKFIYGEDAEPRPRRWLEDTVDLSNS-IEAPYGHQASFSIGPRACGWRFAVEM 505
Db 400 VLHRDPQLWPE-PEEPRPERFSKKNKDSINPYVLLPFG-----TGPNCIGMRFAIMNM 452
Qy 506 KAFLEVTLRVQVEPIISHPEYEHITLIISRPRIVGREK 544
Db 453 KLAIVRLQNFSEKPC-----KETQIFLKNISQGLIRPEK 487
RESULT 4
CP3C CANFA
ID CP3C CANFA STANDARD; PRT; 503 AA.
AC P24463; 1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A12 (EC 1.14.14.1) (CYP3A12) (P450-PBD-1).
GN CYP3A12.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RC MEDLINE=91159488; PubMed=2001406;
RA Ciaccio P.J., Graves P.E., Bourque D.P., Glimsman-Gibson B.,
RA Halpert J.R.;
RT "cDNA and deduced amino acid sequences of a dog liver cytochrome
RT P-450 of the IIIA gene subfamily."
RL Biochim. Biophys. Acta 1088:319-322(1991).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54915; CAA38687.1; -
CC PIR; S14275; S14275.

```



Db 322 HDVQOKQOE---IDTLLPNKELATYDTLVKMEYLDVMVNETLRLYPIAGRLERVCCKD 378  
 Qy 413 EDFIPLAEVIGRDSGVINEVRIKGTMTVMPLFNINRSFIYCEDAEAFRPRMLEDVT 472  
 Db 379 VD-----INGTIFPKGTIVVMPTVALHRDPQHWT-E-PDFRPRERFSKKK 422  
 Qy 473 DGLNS-IEAPYGHQASFTSGPRACGHRPFAVAEMKAFVTLRVOFE-----PI--- 521  
 Db 423 DMNPIYVHPFG-----AGRNCLGMRFAFALMNKIALVLMQNFSGFKCKETQVPLKLG 476  
 Qy 522 ---ISHPEYEHITLIISRPV 540  
 Db 477 KQGLLOPEKPIVLKVSVDGII 498

RESULT 6  
 CP33\_HUMAN STANDARD; PRT; 503 AA.  
 AC P05184;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A3 (EC 1.14.14.1) (CYP11A3) (HLP).  
 GN CYP3A3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=86259780; PubMed=3460094;  
 RA Molowa D.T., Schuetz E.G., Wrighton S.A., Watkins P.B., Kremers P.,  
 RA Mendez-Picon G., Parker G.A., Guzelian P.S.;  
 RT "Complete cDNA sequence of a cytochrome P-450 inducible by  
 RT glucocorticoids in human liver.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5311-5315(1986).  
 RN [2]  
 RP SEQUENCE OF 1-20.  
 RC TISSUE=Liver;  
 RX MEDLINE=85298342; PubMed=3898085;  
 RA Watkins P.B., Wrighton S.A., Maurel P., Schuetz E.G., Mendez-Picon G.,  
 RA Parker G.A., Guzelian P.S.;  
 RT "Identification of an inducible form of cytochrome P-450 in human  
 RT liver.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6310-6314(1985).  
 CC -I- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC nonoxigenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC -I- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -I- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -I- INDUCTION: By glucocorticoids.  
 CC -I- SIMILARITY: Belongs to the cytochrome P450 family.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

ENBL; D00003; BAA00001.1; -  
 ENBL; M13785; AAA35742.1; -  
 PIR; A29410; A29410.  
 HSSP; P14779; 1JP2.  
 DR Genew; HGNC:2636; CYP3A3.  
 DR GO; GO:0005792; C-microsome; TAS.  
 DR GO; GO:0015034; F-cytochrome P450 activity; TAS.  
 DR InterPro; IPR001128; Cytochrome\_P450.

DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01689; EP450IICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT INIT\_MET 0  
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 503 AA; 57428 MW; 9885DIF72958FC0 CRC64;

Query Match 14.0%; Score 399; DB 1; Length 503;  
 Best Local Similarity 28.4%; Pred. No. 6.4e-16;  
 Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;

Qy 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--PTGNFLDILSARTG-----E 61  
 Db 6 LAMETWLLAVSLVLLYLYGTHSHGLFKKUGIGPFTLPFLGN---ILSYKHGFCMFDMDE 62  
 Qy 62 EHAKYREKYGSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYD-----YKFGMAARVL 113  
 Db 63 CHKKYGVWGF---YDG--QQPVLAITDPMIKLVLVKECVSVFTNREPFGVGMKSAI 117  
 Qy 114 RATGQGVVTAEGEAKHRRIRIMIPSLSAQAQKSMVPIFLEKGMELVDOMEDAAEKDMA 173  
 Db 118 SI-----ADEEKKRLRSLSGPTFTSGKLEKEMVPIIAQYGDVLRNLRRE----- 162  
 Qy 174 VGESAGEKKATRETEGVVDKDWVGRATLDVWALAGPDYKSDSLQNKTNELYVAVGLTD 233  
 Db 163 -----RETGKPVTLKDVFGAYSMVDVITSSFGVNVDSLNPPQD----- 200  
 Qy 234 GPAPTLDSEKAIM--WDFV--PYFRIMK--RRHEIPLTQGL-----AVSRVSG 275  
 Db 201 ---PLVENVKLLRDFLDLPFLSLTFVFPFLIPLEVINICVFPREVTNLRKAVRMKE 257  
 Qy 276 IELMEQKQAVLGSASDAQVDKQVGRDILSLVRANIAANLPESOK--LSDEEVLQAQS 334  
 Db 258 SRLEDTQKRV-----DFLQIMDSHKSKKTESHKALSDELVAQSI 300  
 Qy 335 NILFAGYETSSVLTWMPHRLSEDKAVQDKLRBEICQI---DTDMPTLDLNLALPYLEAFV 392  
 Db 301 IFIFAGYETSSVLSFIMTELATPHDVOQKQLEIDAVLENKAPPTTDTVLQMEYLDMMV 360  
 Qy 393 KESRLRDPSPYANRECKLDEDFIAPFPVIGRDSGVINEVRIKGTMTVMPLFNINRSK 452  
 Db 361 NETLRLFIAMLERVCKDVE-----INGMFIPKGVWVWIPSYALHRDP 405  
 Qy 453 FYGEDAEAEFRERLEEDVTDLSNS--IEAPYGHQASFTSGPRACGHRPFAVAEMKAFV 511  
 Db 406 -KYWTEPEKFLPERFSKKNKONIDPIYITPFG-----SGPRNCIGMRFAFALMNKALIR 458  
 Qy 512 TLRRVQFEP 520  
 Db 459 VLQNFSEKP 467

RESULT 7  
 CP3D\_MOUSE STANDARD; PRT; 503 AA.  
 ID CP3D\_MOUSE  
 AC Q64464;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A13 (EC 1.14.14.1) (CYP11A13).  
 GN CYP3A13 OR CYP3A-13.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ddy; TISSUE=Liver;  
 RX MEDLINE=95101705; PubMed=7803471;

RA Yanagimoto T., Itoh S., Sawada M., Hashimoto H., Kamataki T.;  
 RT "Molecular cloning and functional expression of a mouse cytochrome P-  
 RT 450 (Cyp3a-13): examination of Cyp3a-13 enzyme to activate aflatoxin  
 RT B1 (AFB1).";  
 RL Biochim. Biophys. Acta 1201:405-410(1994).  
 CC -!- FUNCTION: Can activate aflatoxin B1 to a genotoxic product.  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR ENBL; X63023; CAA44754.1; -;  
 DR PIR; S50211; S50211.  
 DR HSP; P14779; LUP2.  
 DR MGD; MGI:88610; Cyp3a13.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01689; EP450L1CYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 503 AA; 57492 MW; 15DDE2606B337FCF CRC64;

Query Match 14.0%; Score 398; DB 1; Length 503;  
 Best Local Similarity 27.0%; Pred. No. 7.3e-16;  
 Matches 142; Conservative 74; Mismatches 179; Indels 130; Gaps 20;  
 63 HAKYREKYSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYD-----YKPGMAARVLR 114  
 65 HKYKXWNG--LYDG--RQVLATDPDIITVLVKECYSTFTNRRRFGPVGILKKAIS 119  
 115 IATGDGVVTAAGBAHRRHRIIPSLSAQVSMVPIFEKGMELVDKMDAAEKDQAV 174  
 120 I-----SENBEWKRIRALLSPTFTSGLKEMFPFI-----NQFTDVLVRNMQ 162  
 175 GESAGEKKATRLTEGVGVKDWGRATLDVMAAGFDYKSDSLQNT----- 221  
 163 G--LGEKPT-----SMKIDFGAYSMVDITATSGVNIIDSLNNPQDPFVEKIKLKF 213  
 222 ---NELYAVVGLTDFAPTLDSPKALM--DFVYFRMTKMRHEIPITQGLAVSRVGI 276  
 214 DIFDPLFLS-VTLFPFLTPVFDALNSLFRPDISFFFT-----SVERKEN 259  
 277 ELMEKQKQAVLGASDAQVKKDQVGRDILSLVRANLANLPESQKLSDEVLQAISNL 336  
 260 RMKEKEKQV-----DFQLMINSQNYKTKESHKALSDVEIVAQSIVF 302  
 337 LFAGYETSSVLTWTFHRLSEDKAVQDKUREICQIDTDP-----TLDELNALPYLEAF 391  
 303 IFAGYETSSALSFALYLLAIHPDVQKLQDE---IDAALPNKAPATYDTLQMEYLDV 359  
 392 VKESLRDPPSPYANRECKLDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRS 451  
 360 VNETRLYPIAGRLERVCKTDV-----INGLFIKGTVMVPTFALHKD 404  
 452 KFIYGDAEERFRPERMLEVDTSLSNIEAPYHQASFISSPRACFGWRFAVEMKAFIV 511  
 405 P-KYWPPEFRFRPERFSKKNQDSIN---PYMY-LPFGSGPRNCIGMRFALINMKVALVR 458

QY 512 TLRRVQPEPI-----ISHPEYEHITLIISRRIVGRE 543  
 Db 459 VLQFTVQPCKEVEIPLKLSKQGLQFPENPLLLKVVSRDETWSDE 503  
 RESULT 8  
 ID CP34 HUMAN STANDARD; PRT; 502 AA.  
 AC P08684; Q16757; Q9UK50;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A4 (EC 1.14.13.67) (Quinine 3-monooxygenase)  
 DE (CYP11A4) (Nifedipine oxidase) (NF-25) (P450-PCN1).  
 GN CYP3A4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_taxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88195781; PubMed=3267210;  
 RA Gonzalez F.J., Schmid B.J., Umeno M., McBride O.W., Hardwick J.P.,  
 RA Meyer U.A., Gelboin H.V., Idle J.R.;  
 RT "Human P450PCN1: sequence, chromosome localization, and direct  
 RT evidence through cDNA expression that P450PCN1 is nifedipine  
 RT oxidase.";  
 RL DNA 7:79-86(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87041402; PubMed=3464943;  
 RA Beaune P.H., Umbenhauer D.R., Bork R.W., Lloyd R.S., Guengerich F.P.;  
 RT "Isolation and sequence determination of a cDNA clone related to  
 RT human cytochrome P-450 nifedipine oxidase.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8064-8068(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89108438; PubMed=2563251;  
 RA Spurr N.K., Gough A.C., Stevenson K., Wolf C.R.;  
 RT "The human cytochrome P450 CYP3 locus: assignment to chromosome 7q22-  
 RT qter.";  
 RL Hum. Genet. 81:171-174(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=89093163; PubMed=2463251;  
 RA Bork R.W., Muto T., Beaune P.H., Srivastava P.K., Lloyd R.S.,  
 RA Guengerich F.P.;  
 RT "Characterization of mRNA species related to human liver cytochrome P-  
 RT 450 nifedipine oxidase and the regulation of catalytic activity.";  
 RL J. Biol. Chem. 264:910-919(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=99256082; PubMed=10322772;  
 RA Chen Q., Wu J., Yu Y.;  
 RT "Establishment of transgenic cell line CHL-3A4 and its metabolic  
 RT activation.";  
 RL Zhonghua Yu Fang Yi Xue Za Zhi 32:281-284(1998).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21163842; PubMed=11266076;  
 RA Geilner K., Eiseit R., Huestert E., Arnold H., Koch I., Habert M.,  
 RA Beglmann C.J., Burk O., Bunterfuss D., Becher S., Bishop C.,  
 RA Koebe H.-G., Brinkmann U., Klenk H.-P., Klein K., Meyer U.A.,  
 RA Wojnowski L.;  
 RT "Genomic organization of the human CYP3A locus: identification of a  
 RT new, inducible CYP3A gene.";  
 RL Pharmacogenetics 11:111-121(2001).  
 RN [7]  
 RP SEQUENCE FROM N.A., AND VARIANTS VAL-117 AND ARG-217.  
 RX MEDLINE=21098888; PubMed=11181494;  
 RA Hsieh K.-P., Lin Y.-Y., Cheng C.-L., Lai M.-L., Lin M.-S.,

RA Siest J.-P., Huang J.-D.;  
 RT "Novel mutations of CYP3A4 in Chinese."  
 RL Drug Metab. Dispos. 29:268-273 (2001).  
 RN [8]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RA Zhuge J., Qian Y., Xie H., Yu Y.;  
 RT "Subsequence of a new human cytochrome P450-3A4 cDNA."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP CHARACTERIZATION.  
 RX MEDLINE=97243737; PubMed=9088578;  
 RA Zhang H., Coville P.F., Walker R.J., Miners J.O., Birkett D.J.,  
 RA Wankmcluk S.;  
 RT "Evidence for involvement of human CYP3A in the 3-hydroxylation of  
 RT quinine."  
 RL Br. J. Clin. Pharmacol. 43:245-252 (1997).  
 RN [10]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98122830; PubMed=9456308;  
 RA Zhao X.J., Kawashiro T., Ishizaki T.;  
 RT "Mutual inhibition between quinine and etoposide by human liver  
 RT microsomes. Evidence for cytochrome P4503A4 involvement in their  
 RT major metabolic pathways."  
 RL Drug Metab. Dispos. 26:188-191 (1998).  
 RN [11]  
 RP VARIANTS PRO-221 AND THR-444.  
 RX MEDLINE=20132287; PubMed=10668853;  
 RA Sata P., Sapone A., Elizondo G., Stocker P., Miller V.P., Zheng W.,  
 RA Rao H., Crespi C.L., Gonzalez F.J.;  
 RT "CYP3A4 allelic variants with amino acid substitutions in exons 7 and  
 RT 12: evidence for an allelic variant with altered catalytic activity."  
 RL Clin. Pharmacol. Ther. 67:48-56 (2000).  
 RN [12]  
 RP VARIANTS GLN-161; SER-188; PRO-292; THR-444 AND SER-466, AND  
 RP CHARACTERIZATION OF THE VARIANTS  
 RX MEDLINE=21571796; PubMed=11714865;  
 RA Dai D., Tang J., Rose R., Hodgson E., Bientstock R.J.,  
 RA Mohrenweiser H.W., Goldstein J.A.;  
 RT "Identification of variants of CYP3A4 and characterization of their  
 RT abilities to metabolize testosterone and chlorpyrifos."  
 RL J. Pharmacol. Exp. Ther. 299:825-831 (2001).  
 RN [13]  
 RP VARIANTS ASP-55; GLN-129; ILE-169; HIS-173; MET-362; PHE-372 AND  
 RP LEU-415.  
 RX MEDLINE=21364010; PubMed=11470997;  
 RA Eiselt R., Domanski T.L., Zibat A., Mueller R., Presecan-Siedel E.,  
 RA Huster E., Zanger U.M., Brockmoller J., Klenk H.-P., Meyer U.A.,  
 RA Khan K.K., He Y.-A., Halpert J.R., Wojnowski L.;  
 RT "Identification and functional characterization of eight CYP3A4  
 RT protein variants."  
 RL Pharmacogenetics 11:447-458 (2001).  
 RN [14]  
 RP VARIANTS PRO-14; GLN-161; HIS-173; SER-184 AND PHE-372.  
 RX MEDLINE=21864884; PubMed=11875366;  
 RA Lamba J.K., Lin Y.S., Thummel K., Daly A., Watkins P.B., Strom S.,  
 RA Zhang J., Schuetz E.G.;  
 RT "Common allelic variants of cytochrome P4503A4 and their prevalence in  
 RT different populations."  
 RL Pharmacogenetics 12:121-132 (2002).  
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It performs a variety  
 CC of oxidation reactions (e.g. caffeine 8-oxidation, omeprazole  
 CC sulfoxidation, midazolam 1-hydroxylation and midazolam 4-  
 CC hydroxylation) of structurally unrelated compounds, including  
 CC steroids, fatty acids, and xenobiotics. The enzyme also  
 CC hydroxylates etoposide.  
 CC -!- CATALYTIC ACTIVITY: Quinine + NADPH + O(2) = 3-hydroxyquinine +  
 CC NADP(+) + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- TISSUE SPECIFICITY: Expressed in prostate and liver.  
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other

CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;  
 CC NOTE=CYP3A4 alleles;  
 CC WWW="http://www.imm.ki.se/cypalleles/cyp3a4.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M18907; AAA35745.1; -;  
 CC EMBL; M14096; AAA35744.1; -;  
 CC EMBL; X12387; AAA30944.1; -;  
 CC EMBL; J04449; AAA35747.1; -;  
 CC EMBL; AF182273; AAF13598.1; -;  
 CC EMBL; AF280107; AAG32290.1; -;  
 CC EMBL; AF209389; AAF21034.1; -;  
 CC PIR; A29815; A29815.  
 CC HSSP; P14779; IUPZ.  
 CC Genew; HGNC:2637; CYP3A4.  
 CC MIM; 124010; -;  
 CC GO; GO:0015034; F:cytochrome P450 activity; TAS.  
 CC GO; GO:0004497; F:monooxygenase activity; TAS.  
 CC GO; GO:0006629; P:lipid metabolism; TAS.  
 CC InterPro; IPR001128; Cytochrome P450.  
 CC InterPro; IPR008072; EP450\_CYP3A.  
 CC Pfam; PF00067; P450; 1.  
 CC PRINTS; PR01689; EP450IICYP3A.  
 CC PRINTS; PR03385; P450.  
 CC PROSITE; PS00086; CYTOCHROME P450; 1.  
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 CC Microsome; NADP; Endoplasmic reticulum; Polymorphism.  
 CC INIT MET 0 BY SIMILARITY  
 CC METAL 441 441 IRON (HEME AXIAL LIGAND) (BY SIMILARITY)  
 CC VARIANT 14 14 L -> P (in allele CYP3A4\*14).  
 CC VARIANT 55 55 G -> D (in allele CYP3A4\*7).  
 CC VARIANT 117 117 I -> V (in allele CYP3A4\*4).  
 CC VARIANT 129 129 R -> Q (in allele CYP3A4\*8).  
 CC VARIANT 161 161 R -> Q (in allele CYP3A4\*15).  
 CC VARIANT 169 169 V -> I (in allele CYP3A4\*9).  
 CC VARIANT 173 173 D -> H (in allele CYP3A4\*10).  
 CC VARIANT 184 184 T -> S (in allele CYP3A4\*16).  
 CC VARIANT 188 188 F -> S (in allele CYP3A4\*17; exhibits  
 CC lower turnover numbers for testosterone  
 CC and chlorpyrifos).  
 CC VARIANT 217 217 P -> R (in allele CYP3A4\*5).  
 CC FT FT FTId=VAR\_011605.  
 CC  
 CC Query Match 13.9%; Score 396.5; DB 1; Length 502;  
 CC Best Local Similarity 28.7%; Pred. No. 8.9e-16;  
 CC Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;  
 CC  
 CC QY 13 LAAPSW--ASTAFSLVLAPEERS-SLYNLOG-PNHINY-PTGNFLDILSARTG-----E 61  
 CC Db 6 LAMETVLLVALLVLLVLYGTHSHGULFKLGIPTPLPLPLGN---ILSHKGFQMFME 62  
 CC QY 62 EHAKYRKYGSTLRFAGIAGAPVLINSTDPKVFNVHM-KEAYDY---PKFGMAARVLRIAT 117  
 CC Db 63 CHKYKRWKGF---YDG--QQPVLAITDPDMIKTLVKCYSVFTNRRPFGVGMKSA- 116

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; SS3047; AAB24952.1; -  
 DR PIR; S28168; S28168.  
 DR HSSP; P14779; 1JPZ.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR PRINTS; PF00667; P450; 1.  
 DR PRINTS; PF01689; EP450IICYP3A.  
 DR PRINTS; PF03085; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR OXIDOREDUCTASE; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 503 AA; 57511 MW; D701B6FE83AC8BFB CRC64;  
 Query Match 13.9%; Score 395.5; DB 1; Length 503;  
 Best Local Similarity 28.6%; Pred. No. 1e-15;  
 Matches 155; Conservative 85; Mismatches 187; Indels 115; Gaps 26;  
 QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--PTGNFLDLISARTG-----E 61  
 Db 7 LAVETWLLAVTLVLLYLYGTHSHGLPKLIGPPTPLPLGN---ILSYKKGWTFDME 63  
 QY 62 EHAKYREKYGSTRFAGIAGAPVLNSTDPKVNHYM--KEAYD-----YPPKPMGAARVL 113  
 Db 64 CYKYGKWWGF--YDG--RQPVLAITDPNNIKTVLKECVSVFTNRPPFGVPMKNAI 118  
 QY 114 RIATGCVVTAEGEAHKEHRRIMIPSLSAQAVKSVVPILEKGMELVDKMDAAEKDMA 173  
 Db 119 SI-----AEDEWKRIIRSLSPFTSGKLKEWPIIAKYGDVLRNLREA----- 164  
 QY 174 VGESAGEKATRLT--EGVDVQWVGRATLDVMAAGDYKSDSLQNKTNELXYVAVGLT 232  
 Db 165 -----ETGKPVLLKDVFGAYSDVITSTSGVNIDSLNPPQ----- 201  
 QY 233 DGFAPTLDSFKAIM--WDFV--PYFRTMK--RRHEILPTQGLAVS---RVGIELMEQKQAV 286  
 Db 202 ----PPVENTKLLRDFDLDPFSLISITIPFIIPILEVINISIPFREVTSFLRSVKRI- 256  
 QY 287 LGSASQAVKDKVOGR--DILSLIVRANIAANLPESOKLSDEEVLQISNLLPAGYETS 344  
 Db 257 -----KESRLKDTQKGRVDFLQLMIDSONSKYESHKALSDELVAQSIIIFAGYETT 310  
 QY 345 STVLTWVHRLSEDKAVQDKLREBICQIDTDM-----PTDELNALPYLEAFVKESLRLD 399  
 Db 311 SSVLSPIIYELATHPDVQOKLOE---IDTVLPNKAPETYDTVLQMEYLDVMVNETLRI 367  
 QY 400 PRSEYANRECLKDEDIFLAEPVIGRDSGVINEVRITKGTVMVLPFLFNINRSKIYGEDA 459  
 Db 368 PIAMRLERYCKDVE-----INGIFPKGVVVVMIPIPSYALHDDP--KYWPEP 411  
 QY 460 BEFRPERWLEDVTDLSNS--IEAPYGHQASFTSGPRACFGWRFVAEMKAFLEVTILRRVQF 518  
 Db 412 EKELPERFSKQNDNIDPIYTPFG-----SGPRNCIGMRFALMNMKLAIRVLQNFSE 465  
 QY 519 EP 520  
 Db 466 KP 467  
 RESULT 10  
 CP38\_MOUSE  
 ID CP38\_MOUSE STANDARD; PRT; 504 AA.  
 AC Q84459;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cytochrome P450 3A11 (EC 1.14.14.1) (CYP11A1) (P-450IITAM1) (P-450UT).  
 GN CYP3A11 OR CYP3A-11.

QY 118 GDGVVTAAGBAHRRIRIMIPSLSAQAVKSVVPILEKGMELVDKMDAAEKDMVAGES 177  
 Db 117 ---ISTAEDSEWRRLSLLSPFTSGKLKEMVPIIAQYGDVLRNLREA----- 163  
 QY 178 AGEKATRLT--EGVDVQWVGRATLDVMAAGDYKSDSLQNKTNELXYVAVGLTDGFA 236.  
 Db 164 -----ETGKPVLLKDVFGAYSDVITSTSGVNIDSLNPPQ----- 200  
 QY 237 PTLDSFKAIM--WDFV--PYFRTMK--RRHEILPTQGLAVS---SRVGIEMEQKQAVLQSA 290  
 Db 201 PFVENTKLLRDFDLDPFSLISITIPFIIPILEVINISIPFREVTFNLRKSVKR-----M 255  
 QY 291 SDQAVKDKVOGRDILSLIVRANIAANLPESOKLSDEEVLQISNLLPAGYETSIVLTW 350  
 Db 256 KESRELEDTQKGRVDFLQLMIDSONSKYESHKALSDELVAQSIIIFAGYETTSVLSF 315  
 QY 351 MFRHLSDEKAVQDKLREBICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANRE 408  
 Db 316 IMVELATHPDVQOKLOEIDAFLPNKAPPTYDTVLQMEYLDVMVNETLRLPIAMRLERV 375  
 QY 409 CLKDEDEIPLAEPVIGRDSGVINEVRITKGTVMVLPFLFNINRSKIYGEDAEERPERWL 468  
 Db 376 CKXKDE-----INGMIFPKGVVVVMIPIPSYALHDDP--KYTEPERFLPERFS 419  
 QY 469 EDVTDLSNS--IEAPYGHQASFTSGPRACFGWRFVAEMKAFLEVTILRRVQPEP 520  
 Db 420 KQNKNDIDPIYTPFG-----SGPRNCIGMRFALMNMKLAIRVLQNFSPKPE 466  
 RESULT 9  
 CP38\_MACFA  
 ID CP38\_MACFA STANDARD; PRT; 503 AA.  
 AC P33268; P25231;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A8 (EC 1.14.14.1) (CYP11A8) (P450-MKNF2) (P-450-MK2).  
 GN CYP3A8.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93129612; PubMed=1282830;  
 RA Komori M., Kikuchi O., Sakuma T., Funaki J., Kitada M., Kamataki T.;  
 RT "Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity  
 of the primary sequences to human cytochromes P-450.";  
 RL Biochim. Biophys. Acta 1171:141-146(1992).  
 RN [2]  
 RP SEQUENCE OF 1-22.  
 RC TISSUE=Liver;  
 RX MEDLINE=93287352; PubMed=2500151;  
 RA Ohta K., Kitada M., Hashizume T., Komori M., Ohi H., Kamataki T.;  
 RT "Purification of cytochrome P-450 from polychlorinated biphenyl-  
 treated crab-eating monkeys: high homology to a form of human  
 cytochrome P-450.";  
 RL Biochim. Biophys. Acta 996:142-145(1989).  
 CC -!- FUNCTION: Catalyzes nifedipine and nilvadipine oxidations.  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- INDUCTION: By polychlorinated biphenyl (PCB).  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way

QY	7	LTGALGAAPSWASIAFSLY-LAPARSSLYNQ-----PNHTNYPTGNFLDILSAR 58
DB	3	LVSALSLETWLLAISLVLYRYGTRKHFLFKQGIPIGPKPLP-----FLGTVLNYYKGL 57
QY	59	TGEHAKYREKYGSTL-RFAGIAGAPVLNSTDRKPVNHYM-KEAY-----DYKPGWA 109
DB	58	WKFDMECYK-KYGTWGLPFG--QTPLLAATDDETINKVLVKECFVSFTNRDRDPGVGM 114
QY	110	ARVLRIATGDSGWTAEAGEAHRHRIIMPISLAQAVKSVPIFLEKGMELVDMMEDAAE 169
DB	115	SKAISISKDD-----EMKRYEALLSPFTSGKLKEMFPVIEYQG-DILVKYLQKAK 165
QY	170	KDMAVGESAGKKATRETEGYDVQWVGATLDVMAAGDFYKSDSLONKTNELVAFV 229
DB	166	KG-----KPVTKMDVLGAYSDMVTITSTSGVNDLSNNPDPFVEKAK 208
QY	230	GLT--DGFAPTLDSFKAIMDFV-PYFRTMKRRHEIPLTOGLAVSRVVGIELMEQKQAV 286
DB	209	KLLRDFDFDPLL--FSVVLPPFLPVYEML-----NICMFPKDSIEFPKK---F 252
QY	287	LSASDAQVKKQVQGRDILSLVRA-NIAANUPESQKLSDEVLQAISNLLPAGVETSS 345
DB	253	VDRMKESHLDSQKHRVDFLOLMNSHNNSKDKVSHKALSDMEITQAQSIPIFAGVETTS 312
QY	346	TVLTWFMHRLSEDKAVQDKLREEICQI--BTDMPITDELNALPYLFAFVKESURLDPSP 403
DB	313	STLSFTLHSLATHPDIOKKLQDEIDALEPNKAPPTYDTVMEMEYLDVNLNETURLYPIAN 372
QY	404	YANRECLKDEDFIPLAEPVIGDSGVINEVRIKTGMTVMPLPLNINRSKFIVGEDAEERF 463
DB	373	RLRVCVKDVE-----LNGVYIPKGSVTWMIPTPSVALHHDQHWSE-PEEFQ 416
QY	464	PERWLEDTVDSLNS-IEAPYGHQASFISGRACKGFWFAEVAEMKAFLVTLRRVQPEPII 522
DB	417	PERFSKENKSGSDPVVLPFG-----NGPRNCLGMFALNMKMLATKIMQNFSPQC- 469
QY	523	SUPEYEHITLLISRPRIIVGREK 544
DB	470	--KETQPLKLSRQGLLOPEK-488
RESULT 11		
C341_MOUSE		
ID	C341_MOUSE	STANDARD; PRT; 504 AA.
AC	QJUMAY; 2003	(Rel. 41, Created)
DT	28-FEB-2003	(Rel. 41, Last sequence update)
DT	28-FEB-2003	(Rel. 41, Last annotation update)
DB	Cytochrome P450 3A41	(EC 1.14.14.1).
GN	CYP3A41	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI	Taxid=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ddy; TISSUE=Liver;	
RA	MEDLINE=20239668; PubMed=10775455;	
RA	Sakuma T., Takai M., Endo Y., Kuroiwa M., Ohara A., Jarukamjorn K.,	
RA	Honna R., Nemoto N.;	
RT	"A novel female-specific member of the CYP3A gene subfamily in the	
RT	mouse liver.";	
RL	Arch. Biochem. Biophys. 377:153-162(2000).	
CC	-!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +	
CC	oxidized flavoprotein + H(2)O.	
CC	-!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum	
CC	(Potential).	
CC	-!- TISSUE SPECIFICITY: Expressed in liver. Also expressed in the	
CC	kidneys of female mice, with traces in the stomach, ovary, and	
CC	heart of female mice and in the testis of male mice.	
CC	-!- DEVELOPMENTAL STAGE: Detected immediately after birth in the	
CC	livers of animals of both sexes, but increased with age in	
CC	females, whereas it was gradually reduced in males, resulting in	

**DEVELOPMENTAL STAGE:** Detected immediately after birth in the heart of female mice and in the testis of male mice.



CC predominantly female-specific expression in livers.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB033414; BAA95951.1; --  
 DR MGD; MG1:1858451; CYP3A41.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1  
 DR PRINTS; PR01689; EP450IICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 504 AA; 57959 MW; 2BC645B6B9CD48A CRC64;  
 Query Match 13.7%; Score 391; DB 1; Length 504;  
 Best Local Similarity 26.2%; Pred. No. 1.8e-15;  
 Matches 148; Conservative 100; Mismatches 212; Indels 104; Gaps 23;  
 QY 7 LTGALGLAFAFWASIAFFSLY-LAPRRSSLYNLCQ-PNHT-----NYFTGNFLDIL 55  
 Db 3 LFSALSLEDTWVLLAILLVLLRYGRTGFLFKQIQIPQTPPLPLGTVLVNYKGLW---- 58  
 QY 56 SARTGEHAKYREKYGSTL-RFAGIAGAPVLNSTDPKVNHM-KEAY-----DYPKP 106  
 Db 59 -----KFDMECKYKGTWGLFDG-QMPLFITDPEMKVNLKVECFSVFTNRREFGPV 111  
 QY 107 GMAARVLRATGCVTVASGEAKHRRIMVLSAQAVKSVPIFLKGMELDKWMD 166  
 Db 112 GINSKAIS-----SKDEWKRYALLSFTFTSGKLEKMPVIEGYGDLVKYLMQE 163  
 QY 167 AAEKDMVAGESAGEKATRLTEGVVDKDWVGRATLDVNLALAGFDYKSDSLQNKTNELV 226  
 Db 164 -AEKG-----KPVTKDVLGAYSIDVITSTSGVNVDSLNNPDPFVE 205  
 QY 227 AFVGL--TDGPATLDSFAIMWDFV-PFRFMKRHRHPIPTQGLAVSRVGIELMQKK 283  
 Db 206 KAKGILRVDFDPLV--FSWLFPPFLTPVYEML-----NTCMPFKDSIEFFKK-- 251  
 QY 284 QAVLGASDAQVDKQVQGRDILSLVRANIANLPESOK-LSDEEVLQAISNLLFAGYE 342  
 Db 252 --FVNRKESRLDSKQHRVDFQLMMNHNNSKDKSHKALSDMEITAQSIIVFIFAGYE 309  
 QY 343 TSSTVLTWPHRLSEDAKQVDKLEBEICQI--DTDMPLDELNALPYLEAFVKSLRLDP 400  
 Db 310 TTSSTLSFTLYCIATHPDIOKKLQEBIDETLPNKAPPTVDYTMEMEYLDWNLNLTSLRP 369  
 QY 401 PSPVANRECKLDDFIPLAEPVIGROGVSINVRITKGTVMPLFNINRSKFYIGEDAE 460  
 Db 370 IGRLERFCKQVE-----LNGVYIPKSGTVMIPSYALHHPQHWPE-PE 413  
 QY 461 EFPRLWLDVDTSLNSIEAPYGHQASFTSGRACFGWRFAVAENKALFVTLRVQPEP 520  
 Db 414 EFQPERFSKENKGSID----PYLYM-PFGIGPRNCIGNRFAFMVTKLTKLYMNFSPQ 468  
 QY 521 IISHPEVEHITLISPRVIGREK 544  
 Db 469 C-----QETQIPLKSLRQGLLQPEK 488  
 RESULT 12  
 CP37 HUMAN  
 ID CP37 HUMAN  
 AC P24462;

DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A7 (EC 1.14.14.1) (CYP11A7) (P450-HFLA).  
 GN CYP3A7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=89255154; PubMed=2722762;  
 RA Komori M., Nishio K., Ohi H., Kitada M., Kamataki T.;  
 RT "Molecular cloning and sequence analysis of cDNA containing the  
 RT entire coding region for human fetal liver cytochrome P-450";  
 RL J. Biochem. 105:161-163(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21163842; PubMed=11266076;  
 RA Gellner K., Bisset R., Huestert E., Arnold H., Koch I., Haberl M.,  
 RA Deglmann C.J., Burk O., Buntetuss D., Escher S., Bishop C.,  
 RA Koebel H.-G., Brinkmann U., Klein H.-P., Klein K., Meyer U.A.,  
 RA Wojnowski L.;  
 RT "Genomic organization of the human CYP3A locus: identification of a  
 RT new, inducible CYP3A gene";  
 RL Pharmacogenetics 11:111-121(2001).  
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -1- INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -1- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;  
 CC NOTE=CYP3A7 alleles;  
 CC WWW="http://www.imm.ki.se/cypalleles/cyp3a7.htm".  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D00408; BAA00310.1; --  
 DR EMBL; AF280107; AAG32289.1; --  
 DR FTR; JX0062; JX0062.  
 DR RSP; P44779; IJFZ.  
 DR Genew; HGNC:2640; CYP3A7.  
 DR MIM; 605340; --  
 DR GO; GO:0015034; F:cytochrome P450 activity; TAS.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1  
 DR PRINTS; PR01689; EP450IICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 503 AA; 57470 MW; 087CCEED9BAC314C CRC64;  
 Query Match 13.5%; Score 385.5; DB 1; Length 503;  
 Best Local Similarity 29.0%; Pred. No. 3.8e-15;  
 Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

QY 13 LAAPSW--ASTAFSLYL-APRSSLYNLOG-PNHVNY-FTGNFLDILGARTG-----E 61  
 Db 7 LAVETWLLAVSLILLYLGRTHGLFKLIGPGLPLFLGN---ALSFRKYWTFDWE 63  
 QY 62 BHAKYREKYSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYD-----YKPPGMAARVL 113  
 Db 64 CYKIRKVG-----YDCCQPMALAITDPMIKTVLVKCYSVFTNRFPFGVFWKNAI 118  
 QY 114 RIATGDGVVTAAGBAHRRIRIMPSIAQVSMVPIFEKGMELVDKMDAEAKOMA 173  
 Db 119 SI-----AEDBEWKIRISLSPFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164  
 QY 174 VGESAGEKATRLT-EGVDVKWVGGRATLDVVALAGFDYKPSLONKTNELVAFVGLT 232  
 Db 165 -----BTGRPVTLKHVFGAYSMDVTSTSGVSDLSNPNOD-----PFVENT 207  
 QY 233 D---GFAPTLDSFKRAIMWDFVYPRTRMKRHEIPLTQGLAVS---RRVGIELMEQKKQAV 286  
 Db 208 KLLRFNP-LDPFVLSIKVF-PFL-----TPILEALNITVPRKVISFLTYSVKQIK 257  
 QY 287 LGSASDAQVKKDQGR--DILSLVZVANTAANLPESQKLSDEVLQAISNLLFAGYETS 344  
 Db 258 EGRL-----KETQKRVDFLQMLDSQKSETHKALSDLELWQASIIIFAGYET 310  
 QY 345 STVLTMPHRLSEDKAVQDKLREICQIDTDM-----PTLDELNALPYLEAFVKESLRLD 399  
 Db 311 SSVLSFIIYELATHPDVQKQVKE---IDTVLPNKAPTVDYTVLQLEYLDMVNVNLTRLF 367  
 QY 400 PPSVYANRECLKDDFTPLABPVIRGDSVINEVRIKGTWMLPLFNINRSKFIYGEA 459  
 Db 368 PYAMELERVCKDVE-----INGMFIKGVVVMIPSVLHHPD-KYWTGP 411  
 QY 460 BEFRPERLEVDVSLNS-IEAPYHQASFTSGPRACFGWRFAVEMKAPFLFVTLRRVQF 518  
 Db 412 EKFLPERFSKKKNIDPIYITPFG-----SGPRNCIGVRPALVNMKLVALVEVLQNFSP 465  
 QY 519 EP 520  
 Db 466 KP 467  
 RESULT 13  
 CP35 HUMAN STANDARD; PRT; 502 AA.  
 AC 01-PEB-1991 (Rel. 17, Created)  
 DT 01-PEB-1991 (Rel. 17, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Cytochrome P450 3A5 (EC 1.14.14.1) (CYP11A5) (P450-PCN3).  
 GN CYP3A5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]\_SEQUENCE FROM N.A.  
 RP MEDLINE=89278095; PubMed=2732228;  
 RA Aoyama T., Yamano S., Waxman D.J., Lapenson D.P., Meyer U.A.,  
 RA Fischer V., Tyndale R., Inaba T., Kallow W., Gelboin H.V.,  
 RA Gonzalez F.J.;  
 RT "Cytochrome P-450 hPCN3, a novel cytochrome P-450 IIA gene product  
 RT that is differentially expressed in adult human liver. cDNA and  
 RT deduced amino acid sequence and distinct specificities of cDNA-  
 RT expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and  
 RT cyclosporine.";  
 RL J. Biol. Chem. 264:10388-10395 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Sapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP SEQUENCE OF 1-106 FROM N.A.  
 RX MEDLINE=21163842; PubMed=11266076;  
 RA Gellner K., Eisele R., Hustert E., Arnold H., Koch I., Haberl M.,  
 RA Deglmann C.J., Burk O., Buntefuss D., Escher S., Bishop C.,  
 RA Kosbe H.-G., Brinkmann U., Klenk H.-P., Kleine K., Meyer U.A.,  
 RA Wojnowski L.;  
 RT "Genomic organization of the human CYP3A locus: identification of a  
 RT new, inducible CYP3A gene.";  
 RL Pharmacogenetics 11:111-121 (2001).  
 RN [4]  
 RP SEQUENCE OF 1-24 FROM N.A.  
 RX MEDLINE=96152836; PubMed=8569713;  
 RA Schuetz J.D., Schuetz E.G., Thottassery J.V., Guzelian P.S., Strom S.,  
 RA Sun D.;  
 RT "Identification of a novel dexamethasone responsive enhancer in the  
 RT human CYP3A5 gene and its activation in human and rat liver cells.";  
 RL Mol. Pharmacol. 49:63-72 (1996).  
 RN [5]  
 RP VARIANT CYP3A5\*2.  
 RX MEDLINE=96192071; PubMed=8619878;  
 RA Jounaidi Y., Hyafilles V., Gervot L., Maurel P.;  
 RT "Detection of CYP3A5 allelic variant: a candidate for the polymorphic  
 RT expression of the protein?";  
 RL Biochem. Biophys. Res. Commun. 221:466-470 (1996).  
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -1- INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -1- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;  
 CC NOTE=CYP3A5 alleles;  
 CC WWW="http://www.imm.ki.se/CYPalleles/cyp3a5.htm".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC EMBL; J04813; AAA02993.1; --  
 CC EMBL; BC033862; AAA33862.1; --  
 CC EMBL; AF280107; AAG32288.1; --  
 CC EMBL; U35912; AAB00083.1; --  
 CC PIR; A34101; A34101.  
 CC HSP; P14779; 1JPZ.  
 CC Genew; HGNC:2638; CYP3A5.

DR MTM; 605325; --  
 DR GO: 0005792; C: microsome; TAS.  
 DR GO: 0015034; F: cytochrome P450 activity; TAS.  
 DR GO: 0004497; F: monooxygenase activity; TAS.  
 DR GO: 0008202; P: steroid metabolism; TAS.  
 DR InterPro: IP0001128; Cytochrome P450.  
 DR InterPro: IP0008072; EP450\_CYP3A.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR01689; EP450IICYP3A.  
 DR PROSITE: PS00086; CYTOCHROME P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum; Polymorphism.  
 FT METAL 441 441 IRON (HEME AXIAL LIGAND)  
 FT VARIANT 398 398 /FTID=VAR 008385.  
 FT SEQUENCE 502 AA; 57108 MW; D5A2302E2638717 CRC64;  
 Query Match 13.5%; Score 385; DB 1; Length 502;  
 Best Local Similarity 28.1%; Pred. No. 4.1e-15;  
 Matches 151; Conservative 94; Mismatches 196; Indels 106; Gaps 25;  
 QY 13 LAAPSW-ASIAFSLYL-APRRSLYLNQ-Q-PNHTNY-FTGNFLDILSARTG-----E 61  
 Db 7 LAVETWLLAVSLVLLYLYGTRTHGLFKRLGIPGTPPLLGN---VLSYRQGLWKFDTE 63  
 QY 62 EHAKYREKYGSTLRPAGIAGAVLNSTDPKFNHYM-KEAYDPKPGMAARVLRATG-- 118  
 Db 64 CYKYGKMGWT---YEG-QLPVLATDPDVTIRVLVKECY-----SVFTNRSILGPV 111  
 QY 119 ----DGVVTAAGBAKRRHRRIMIPSLAQAVKSMVPIFEKGMELVDKMDAEKQNAV 174  
 Db 112 GFMKSAISLAEDBEWKRIIRSLISPTFTSGKLKEMFPPIAQYGDVLRNLRRE-AEKG-- 167  
 QY 175 GESAGEKKAFLTEGTGVVDKDWGRATLDVMAAGDYKSDSLQ-----KINELYVAF 228  
 Db 168 -----KPTVLIDFGAYSMVITGTSGVNIIDSLNPPQDPFVESIKKFKP 213  
 QY 229 VGLTDGAPATLDSFKAIMDVFVYFRMKRHEIPLTQGLAVS--RRVGIELMEQKQAV 286  
 Db 214 ----GFLDEL-FLSILPFL-----TPVFEALNVSLPDKTINFLSKS----- 252  
 QY 287 LGSASDAVNDKQVQGRDILSLAVRANIANLPESQKLSDEVLQAIINLLFAGVETST 346  
 Db 253 VNRKKSLRNDKQKRLDFLQMLDSQNKETESHKALSDLELAQSIIFIFAGVETTS 312  
 QY 347 VLTWFMHRLSDKAVQDKLREIICQI--DTPDPTLDELNALPYLEAFVKESLRDPPSPY 404  
 Db 313 VLSFTLYELATHPDVQQLQKEIDAVLPNKAPPTYDAVVQMEYLDWVYNETLRLFPVAIR 372  
 QY 405 ANRECLKDDEPIELAEPIVIGRDSVINEVRIKGTWMLPLFNIRSKFEIYGEDAEERFP 464  
 Db 373 LERTCKKDVE-----INGVFIPKGMVVIPTVALHDP-KYWTEPEERFP 416  
 QY 465 ERWLEDVTDLSNS-IEAPYGHQASFTSGPRACGMRFAVAEMKAFILFVTLRVRQEP 520  
 Db 417 ERFSKK-KDSIDPIYITPFG-----TGPENCIGKRFALMNMKLIRVLQNFSEK 466

RESULT 14  
 ID CP3S BOVIN STANDARD; PRT; 507 AA.  
 AC P79102;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 3A28 (EC 1.14.14.1) (CYP11A28).  
 GN CYP3A28.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;

RP SEQUENCE FROM N.A.  
 RC STRAIN=Hereford; TISSUE=Liver;  
 RA Natsuhori M., van Raak M., Ligtnerberg M., Kleij L., ten Berge D.,  
 RA Zweers-Zeilmaier W.M., de Groene E.M., van Miert A.S.J.P.A.,  
 RA Witkamp R.F., Horbach G.J.M.J.;  
 RT "Isolation of a bovine full length cytochrome P450 (CYP3A) cDNA  
 RT sequence and its functional expression in V79 cells.";  
 RL Environ. Toxicol. Pharmacol. 3:17-24(1997).  
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; Y10214; CA471266.1; --  
 DR HSSP; P14779; IJFZ.  
 DR InterPro: IP0001128; Cytochrome P450.  
 DR InterPro: IP0008072; EP450\_CYP3A.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR01689; EP450IICYP3A.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 FT SEQUENCE 507 AA; 58152 MW; FDFD846SCC96DF57 CRC64;  
 Query Match 13.5%; Score 385; DB 1; Length 507;  
 Best Local Similarity 25.5%; Pred. No. 4.1e-15;  
 Matches 140; Conservative 94; Mismatches 192; Indels 122; Gaps 22;  
 QY 38 LQGNHTNYFTGNFLDILSARTGEEHAKYR-----EKYGSTLRPAGIAGAVLNS 87  
 Db 38 IFGPRPVYF-----GSTMAVHKGIPEDFNQCFKYGKQMGFYE-GRQPMIAI 84  
 QY 88 TQPKVFNHYM-KEAYD-----YKPGMAARVLRATGCVVTAEGEAHRRHRRIMIPS 139  
 Db 85 TPDPIIKTVLVEKCVSVTNRRIFGPMGIMKYAISLAW-----DEQWKRITLLSPA 136  
 QY 140 LSAQAVKSMVPIFEKGMELVDKMDAEKMDAVGESAGEKKAATRLTGTGVVDKQWGR 199  
 Db 137 FTSGKLKEMFPPIIQYGDVLRNLRKE-AEKG-----NPNVNMKDMFGA 178  
 QY 200 ATLQWALAGFDYKSDSLQNKTNELNYAVFGLTQGPATLDSF--KAIMMDFV-PYERTM 256  
 Db 179 YSMDEVITGTAFGVNIIDSL-NNPDFVEHSHKNLRLFRP-FDFPILSILLPFLNPVEIL 236  
 QY 257 KRRHEIPLTQGLAVSRVGIEMEQKQAVLGSASDAQVKKQVGGVGRIDISLVANIAA 316  
 Db 237 -----NITLFPKSVDFFTKSVKKI-----KESRLTDKQNRVLDLLQMINSONSK 282  
 QY 317 NLDESOKLSDEVLQAIINLLFAGVETSTVLTWFMHRLSDKAVQDKLREIICQIDTM 376  
 Db 283 EIDNHKALSDIELVAQSTIFFGYETTSLSFTSIYELTTHPHVQQLQEE---IDATP 339  
 QY 377 -----PTLDELNALPYLEAFVKESLRDPPSPYANRECLKDDEFTPLAEPVIGRDSVIN 431



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 38.5 Seconds

(without alignments)  
4564.770 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852  
Sequence: 1 MFILVLTGALGAARFNAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	425	14.9	576	3 Q8J2Q5	Q8J2Q5 gibberella
2	419	14.7	593	10 Q9C6S0	Q9C6S0 arabidopsis
3	419	14.7	595	10 Q93VK5	Q93VK5 arabidopsis
4	417	14.6	503	6 Q97689	Q97689 sus scrofa
5	417	14.6	517	10 Q8LIR5	Q8LIR5 oryza sativ
6	403.5	14.1	503	6 Q95L35	Q95L35 sus scrofa
7	403	14.1	515	10 Q9LUD2	Q9LUD2 arabidopsis
8	401	14.1	515	10 Q9W1E1	Q9W1E1 arabidopsis
9	400.5	14.0	519	10 Q9ASR3	Q9ASR3 arabidopsis
10	397.5	13.9	519	10 Q8LHV0	Q8LHV0 oryza sativ
11	394	13.8	503	4 Q85SK3	Q85SK3 homo sapien
12	393.5	13.8	544	10 Q9FEI1	Q9FEI1 oryza sativ
13	392	13.7	497	11 Q8CJF2	Q8CJF2 rattus norv
14	389	13.6	492	5 Q8ISJ7	Q8ISJ7 mamestra br
15	387	13.6	523	10 Q8LQ77	Q8LQ77 oryza sativ
16	386.5	13.6	511	6 Q8HYE6	Q8HYE6 capra hircu

17	385.5	13.5	535	4 Q9H241	Q9H241 homo sapien
18	384.5	13.5	531	10 Q9PED3	Q9PED3 oryza sativ
19	382.5	13.4	527	10 Q8LGM8	Q8LGM8 zea mays su
20	382	13.4	430	4 Q7Z448	Q7Z448 homo sapien
21	381	13.4	503	6 Q8HZK1	Q8HZK1 canis famil
22	381	13.4	520	10 Q8L4Q4	Q8L4Q4 oryza sativ
23	380.5	13.3	505	10 Q9LUD0	Q9LUD0 arabidopsis
24	374	13.1	533	5 Q964T2	Q964T2 blattella g
25	371	13.0	507	10 Q8L3Q8	Q8L3Q8 oryza sativ
26	368.5	12.9	523	10 Q9FDZ1	Q9FDZ1 oryza sativ
27	365	12.8	512	10 Q9LUC8	Q9LUC8 arabidopsis
28	365	12.8	520	10 Q48786	Q48786 arabidopsis
29	364.5	12.8	503	5 Q8WQM9	Q8WQM9 anopheles g
30	363.5	12.7	511	10 Q64631	Q64631 arabidopsis
31	363.5	12.7	512	10 Q9LUD3	Q9LUD3 arabidopsis
32	363	12.7	502	11 Q06884	Q06884 rattus norv
33	362.5	12.7	523	10 Q9CAD6	Q9CAD6 arabidopsis
34	362	12.7	506	10 Q9LUD1	Q9LUD1 arabidopsis
35	361	12.7	528	10 Q9FE60	Q9FE60 oryza sativ
36	360	12.6	512	10 Q8W2N3	Q8W2N3 vicia sativ
37	359.5	12.6	512	10 Q9LUC9	Q9LUC9 arabidopsis
38	357.5	12.5	512	10 Q9LUC5	Q9LUC5 arabidopsis
39	357.5	12.5	528	10 Q8L174	Q8L174 zea mays su
40	356.5	12.5	523	10 Q8LAQ9	Q8LAQ9 arabidopsis
41	356	12.5	504	11 Q9EQW4	Q9EQW4 mus musculu
42	355.5	12.5	512	10 Q9ZW95	Q9ZW95 arabidopsis
43	354.5	12.4	506	10 Q7XTY4	Q7XTY4 oryza sativ
44	354.5	12.4	560	10 Q8W372	Q8W372 oryza sativ
45	354.5	12.4	560	10 Q7XD91	Q7XD91 oryza sativ

## ALIGNMENTS

### RESULT 1

Q8J2Q5	PRELIMINARY;	PRT;	576 AA.
ID Q8J2Q5			
AC Q8J2Q5;			
DT 01-MAR-2003 (TrEMBLrel. 23, Created)			
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE Fun15p.			
GN Fun15.			
OS Gibberella moniliformis.			
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.			
OX NCBI_TaxID=117187;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=M-3125;			
RX MEDLINE=93343882; PubMed=10413619;			
RA Proctor R.H., Desjardins A.E., Plattner R.D., Hohn T.M.;			
RT "A polyketide synthase gene required for biosynthesis of fumonisin			
RT mycotoxins in Gibberella fujikuroi mating population A.";			
RL Fungal Genet. Biol. 27:100-112(1999).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=M-3125;			
RX MEDLINE=21585628; PubMed=11728154;			
RA Seo J.A., Proctor R.H., Plattner R.D.;			
RT "Characterization of four clustered and coregulated genes associated			
RT with fumonisin biosynthesis in Fusarium verticillioides.";			
RL Fungal Genet. Biol. 34:155-165(2001).			
RN [3]			
RP SEQUENCE FROM N.A.			
RC STRAIN=M-3125;			
RX Proctor R.H., Brown D.W., Plattner R.D., Desjardins A.E.;			
RA "Co-expression of fifteen contiguous genes delineates a fumonisin			
RT biosynthetic gene cluster in Gibberella moniliformis.";			
RL Fungal Genet. Biol. 0:0-0(2003).			
DR EMBL; AF155773; RAN74818.1;			
DR GO; GO:0003676; F:nucleic acid binding; IEA.			
DR GO; GO:0006118; P:electron transport; IEA.			

```
DR InterPro: IPR001128; Cytochrome_P450.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00067; p450.1.
DR PRINTS: PRO0385; P450.
DR PROSITE: PS00030; RRM_RNP_1; 1.
DR SEQUENCE 576 AA; 64918 MW; B1693773CA36A8FB CRC64;

Query Match 14.9%; Score 425; DB 3; Length 576;
Best Local Similarity 24.9%; Pred. No. 1.4e-22;
Matches 147; Conservative 100; Mismatches 230; Indels 114; Gaps 21;

Qy 13 LAAFSWASIAFFSLYLAPRSSLYNQGN--HTNYFTGNFLDILSARTGEEHAKYREKY 70
Db 50 ISHFVNSAI-TWTFPSPLR-----QLPNVSDGHLKETRLVSEPRGVPSQSWINSL 102
Qy 71 GS-----TLRFAGIAGAPVLNSTDPKVNHYM-KEAYDPKPGMAARVLRIATGDDVVTA 125
Db 103 SNRPVDLARYSESLGFERLLIISPKALAEVLTTSYDFRKPGLIVSELKQATGMGVLLAE 162
Qy 126 GEAKHRRHRIWISLSAQAVKSMVPIFLEKGMELVDKMDAAEDMAVGSAGEKKATR 185
Db 163 GSEHKSORKALQAFNVHKNLYPFWDVAGEFATVL-----EKQIFTG-----TPRTS 212
Qy 186 LETGVVDKVGKRALDVMALAGDPYKDSLQNKTNELYAVFVGLTDFGAFPTLDS---- 241
Db 213 DTTAVIDVDWAGKATLIDIGRAGMGQGFDAIQNDSDRLHQAYRMI---FEPGRGAIFLA 269
Qy 242 -----FKALMWDFVFPYFERTMKRHEIPTLTGLAVSRVGVIELMEOKQAVLGSADQAVD 296
Db 270 LLRIIFERLIVLPLRNKRMRH-----GIQVIRSKQELIRKKEI---KROKAGVD 321
Qy 297 KKVQGGRIISLLVRANIAANPESOKLSDEEVLQISNLLFAGYBTSVTLTMFHL 356
Db 322 N---SGNDILTVALNGV-----FTDEQLIDQLMTFLAAGHETTATLTWAIYILC 369
Qy 357 EDKAVQKLBREICQ-----IDTMTPLDELN-----ALPYLEAFVKESLRLLDPSP 403
Db 370 KQPEVQNRLEEREIMHFPNPKWPRSEPSNLTQQAIDFKPLYNVVCLEVMRYFAPIP 429
Qy 404 YANRECKLDEDFIPLAEFVIGRDSVINEVRIKGTWVWLPLFNFNRSFYIGEDAEFR 463
Db 430 LTWEATCD-----TTLHTFVPAQTRIIAPRVNDRSALWGPDANNFN 474
Qy 464 PERMLEVDVTSLSNIEAPYGHQ-----ASFGSPRACFGWRFAVEMKAF 508
Db 475 PDR-----NOKRDSAAPEVTQEVGRTEARSNYADLTFLHGRSCIGQSFAVFEAIL 529
Qy 509 LFTVLRVRQRPILISHPEYHITLIISR---PRIVGREKEGYOMRLQVKPV 556
Db 530 LATIANFEFO-IEDESILDERNISIRGATSIRVG-----GLKYVRVPI 573

RESULT 2
Q9C6S0
ID Q9C6S0 PRELIMINARY; PRT; 593 AA.
AC Q9C6S0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P450, putative.
GN F5M6.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White C., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
```

```
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., L.,
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R.C., Marziali A.,
Millscher J., Miranda M., Nguyen M., Niernan M.C., Osborne B.I.,
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
Sakano H., Salzer S.L., Schwartz J.R., Shim J., Shim P., Southwick A.M.,
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.B.,
Uteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
"Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
Nature 408:816-820(2000).
RL -/- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC ENBL; AC079041; BAG50718.1; -.
DR PIR; F86441; F86441.
DR GO; GO:0004497; F-monooxygenase activity; IEA.
DR GO; GO:0006118; P-electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PRO0385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 593 AA; 66643 MW; E80CB9B8B2BD199 CRC64;

Query Match 14.7%; Score 419; DB 10; Length 593;
Best Local Similarity 26.4%; Pred. No. 4.1e-22;
Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

Qy 50 NFILDSARTG--BEHAKYREKYGT-----LRFAGI-----AGAPVLN 86
Db 93 NVLDFMDFMGSDQDPYKVPKAGSIQVAREAFPIYELFYGGIFRLTFGKSFLLI 152
Qy 87 STDPKVFNHMKV-AYDYPKPGMAARVLRIATGQVTAEGEAKHRRIMPSLSAQAV 145
Db 153 VSDPSIAKHILKDNKAVASK-GILAEILDFVNGKGLIPADGEIWRRRRAIVPALHQYV 211
Qy 146 KSMVPIFLEKGMELVDKMDAAEDMAVGSAGEKKATRLETEGVDVKWVGKATLDVM 205
Db 212 AMISLFGASDRLCQKL--DAA-----ALKGEVEVESLFSRLTLDII 253
Qy 206 ALAGPYKDSLSQNKTNELYAVFVGLTDFGAFPTLDSFKALMWDFVFPYFRTMKRHEIPT 265
Db 254 GKAFVNYDFDSLNTDGTGVIEAVYTVLREADRSVSP--VWD-IPWIKDIS-----PRQ 305
Qy 266 QGLAVSRVRVGIEMEOKQAVLGSADQAVDKQVQ-----GRD--ILSLVRANIAA 316
Db 306 RKVATSLKINDTLD-----LIATCKRVEEELQFHEEYMRERDPSILHFL----- 353
Qy 317 NLPESOKLSDEEVLQISNLLFAGYBTSVTLTMFHLSEDKAVQDKLREICQIDTD- 375
Db 354 -LASGDVSSKQLRDDLTMTLTIAGHETSAVLTWTFTLTTPSPVAKLQBEVSDIGDR 412
Qy 376 MPTLDELNALPYLEAFVKESLRLDPPSPYANRECKLDEDFIPLAEFVIGRDSVINEVRI 435
Db 413 FPIQMKKLVKTTVRVMSLSRLYPQPVLPERSI-----DNDILGEVPI 457
Qy 436 TKGTWMLPLFNINRSKFIYGSDEEFPERWLEVDTD-----SLSIEAPYGHQASFSIG 491
Db 458 KRGEFISVWNLHRSR-LHWDDAEKFNPERPLDGNPNETNQNFSLYPFG-----GG 510
Qy 492 PRACFGWRFAVEMKAFVTLRRVQFE 519
Db 511 PRKICGMFASFNVAIAMLIRRFNFQ 538

RESULT 3
Q93VK5
ID Q93VK5 PRELIMINARY; PRT; 595 AA.
AC Q93VK5;
```

01-DEC-2001 (TrEMBLrel. 19, Created)  
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
ATG31800/68069.m00159.  
Arabidopsis thaliana (Mouse-ear cress).  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
NCBI\_TaxID=3702;  
[1]  
SEQUENCE FROM N.A.  
Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
Ecker J.R.;  
RA "Arabidopsis cDNA clones.";  
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE FROM N.A.  
Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
Ecker J.R.;  
RA "Arabidopsis cDNA clones.";  
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,  
Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,  
Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,  
Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,  
Yamada K., Yu G., Yuan S., Davis R.W., Theologis A., Ecker J.R.;  
RA "Arabidopsis ORF clones.";  
RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AY058173; AAL25587.1; -.  
DR EMBL; AY058446; AAL08302.1; -.  
DR EMBL; AY142017; AAM98281.1; -.  
DR GO; GO:0004497; F:monooxygenase activity; IEA.  
DR GO; GO:006118; P:electron transport; IEA.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME P450; 1.  
DR Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 595 AA; 66845 MW; A548A33A48323B7D CRC64;  
Query Match 14.7%; Score 419; DB 10; Length 595;  
Best Local Similarity 26.4%; Pred. No. 4.1e-22;  
Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;  
QY 50 NFILDSARTG--EEHAKREKYGST-----LRFAGI-----AGAPVLN 86  
DB 95 NVLDFMFDWGTGSDQDPKVPKPEAKSGTQAVRNEAFFIPLYEFLTYGGIFRITFGPKSLI 154  
QY 87 STDPKVFNVHMK-EAYDPKFGMAARVLRIATGDGVTAAEGEAHRRHRIIMPISLAQAV 145  
DB 155 VSDPSTAKHILKNAYASK-GILAEILDFVGMKGLIPADGIWRRRRRAIVPALHQYV 213  
QY 146 KSMVPIFEKMBELVDMMEDAAEKDMVAGESAGEKATRLTEGVDVNDKWDVGRATLDM 205  
DB 214 AAMISLFGESDLSCKL--DAA-----ALKGEVEVESLFSRUTLDII 255

QY	175	GSAGEKATRLTEGVGDVQDWGVRATLDVWALAGFDYKSDSL	-----QKTNELYV 226
Db	168	-----KPVTKMDIFGAYSDVITSTAFQVNTDFLNNPQDPFVENSKKLLKF	213
QY	227	AFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRRHEIPLTOGLAVS	--RRVGIELMEQKK 283
Db	214	SP-----FSPU-----FLSII-----FPFPL-----	TPILEVNTVLPFSKVNFVMSRK 254
QY	284	QAVLGSASDAQVDKDVGRDILSLVRANIANVLPESQKLSDEEVLQISNLLFAGYET	343
Db	255	R-----MKESRLKQKQTHRVDFLOLMINSQNSKETDTHKGLSDEELVAQGVFFI	PAGYET 309
QY	344	SSVLTWTFMHRLSDEKAVODKLRBEICQI	---DTPMTPLDNLNALPYLEAFVKESRLRLDP 401
Db	310	TSSLSLLVIELATHPDVQQKQLQEIIDATPFSKALPSYDALAQEWYLDWMVNEILLYPI	369
QY	402	SPYANRECLKQEDFIPLAEPVIGRGGVINEVRITKGTWMLPLFNINRSKIYGEDABE	461
Db	370	AARLERVCKQVE-----IHGVSYPKGTVMVVFSTHRRDDELMPPE	-PBE 413
QY	462	FPERWLEDVTDLSNLSIEAPYCHQASRTISGRACFGWRFAEAKAFLFVTLRRVQFEP	521
Db	414	FRPERFSKKNKDSIN-----PPTY-LPFGTGPBNCIGMRFALMKNKALVRVLQNSFKPC	468
QY	522	-----ISHPEYEHITLIISRPVIG 541	
Db	469	KETQTPKLSSQGLIOPEKPIILLKVPRDGTGV 501	
RESULT 5			
Q8LIRS PRELIMINARY; PRT; 517 AA.			
ID	Q8LIRS	AC	Q8LIRS;
DT	01-OCT-2002	(TREMELrel. 22, Created)	
DT	01-OCT-2002	(TREMELrel. 22, Last sequence update)	
DT	01-JUN-2003	(TREMELrel. 24, Last annotation update)	
DE	DE	Putative cytochrome P450.	
GN	OUI332	C12.12.	
OS	Oryza sativa	(japonica cultivar-group).	
OC	Eukaryota;	Viridiplantae;	Streptophyta;
OC	Spermatophyta;	Magnoliophyta;	Liliopsida;
OC	Ericaroidae;	Oryzae;	Poaceae;
OX	NCBI_TaxID=39947;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Nipponbare;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC		
RT	clone:OUI332.C12.12."		
RL	Submitted (JUN-2001) to the-EMBL/GenBank/DBSJ databases.		
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; AP003752; BAC10039.1; -.		
DR	Gramene; Q8LIRS; -.		
DR	GO; GO:0004497; F:monoxygenase activity; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome_P450.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR00385; P450.		
DR	PROSITE; PS00086; CYTOCHROME P450; 1.		
DR	Heme; Monoxygenase; Oxidoreductase.		
SK	SEQUENCE 517 AA; 58038 MW; 3C3E357450265138 CRC64;		
Query Match 14.6%; Score 417; DB 10; Length 517;			

Query Match 14.6%; Score 417; DB 10; Length 517;

QY	6	LLTGALGAAFSWASIAFSSLYLAP----	RRSSLYNQGNHNYFTGNEFDLSARTGE	61
		:::::	:::::	
DB	7	MYAAAAVAASWAPDAVVKLVMPRAITR	LRLAQGVGGEGY-RPFSGNLGEI--RRLRD	63
		:::::	:::::	
QY	62	EHA-----	KYREKYGSTLRFAGIAGAPVLNSTD	101
		:::::	:::::	
DB	64	EGAGVVLVDVSHDFPIVOPHFKEWIPLY	GKTFWY-WFGARPTICLADVSMRQVLS	122
		:::::	:::::	

QY	102	D	-YPK----	PGMAARVLR	IATGDGVTWTAEGBAHGRHRIMTPSLSAQAQVKSMPFLFLEKGM	15
Db	123	GMYPNVSNPYFABLL-----	GKLVLVDGDEWRHKRVVHPAFNMOKLKKMT-----	:	:	170
QY	158	ELVDRQMEDAAEKDMAYGESAGEKKA	KATRLTEGVDDVKDWVGR-----	ATLDVMALAGF--D	211	
Db	171	-----VTNSDCAQSWISWE\$-----	ELGTKGDIVEILSRFEELTADVISHTAFGSS	219		
QY	212	YKSD-----	SLOKNTNELIYAVFGLTGDFAPTDLDSFKAIMWDVPVFRPTWKRRHEIP	263		
Db	220	YKEGKQVFACRELQ-----	FLAF-----STFLSI-----	: :	247	
QY	264	LTOGLAVERRVGIEMBOKKQAVLGASDAQDKOV-	OGRDLSLLVRANIAANLPES	321		
Db	248	GSYUFTKNLUKTSVDKVRSMUTDI	IKSELNNKVAGYGNLLGLMLEA-CAPEHGES	306		
QY	322	Q-KLSDEEVLQAIGNLLFAGYETSTSVLTWMFHRLS	EDKAVQKLBREIC-QIDTDMPTL	379		
Db	307	QPQL\$MDEIIAECKTFFTFAGHDTTSHLLTWMTFELL	THPEWQEKLREBATECDGKVPTG	366		
QY	380	DELNALPVLEAFVKE\$URLDPSPSYANRECLKOEDFF	TPLAEPVIGRDG\$VINEVRITKT	439		
Db	367	DMNLKTLKVNMFLELTURLYGVAFIORRVNAELE-----	LGGITVPEGT	411		
QY	440	WYMLPFLFNINSKFITYGEDASEPPEPERKAWLEDVDT	SLNSIBAPYCHQCAFISGPRACFGMR	499		
Db	412	VLSIFPIATHRDKKEVGEDADIFXERFKNGSKAGYPNA---	LLSPSGSGPRACIGQN	467		
QY	500	FAVAMKAPLFVTVIRRVQFEPFIISHPEYEHM--	TLIISRPR	538		
Db	468	FAMIEAKAVIAMILQRFSF---TLSPKYVHVPTDVITLRPK	505			
RESULT 6						
ID	Q95L35	PRELIMINARY;		PRT;	503 AA.	
QC	A95L35;					
DT	01-DEC-2001 (TrEMBLrel. 19, Created).					
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)					
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)					
DE	Cytochrome P450 3A.					
GN	CYP3A.					
OS	Sus scrofa (Pig).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.					
NCBI_TaxID=9823;	[1]					
RN	SEQUENCE FROM N.A.					
RP	Soucek P., Zuber R., Anzenbacherova E., Linka M., Anzenbacher P.;					
RT	"CDNA of minipig cytochrome P450 3A.";					
RL	Submitted (SEP-2001) to the ENBL/GenBank/DBJ databases.					
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.					
DR	ENEL, A8424780; AALL3316.1; ...					
DR	GO: GO:0004497; Fmonooxygenase activity; IEA.					
DR	GO: GO:0016712; Foxidoreductase activity, acting on paired d. . ; IEA.					
DR	GO: GO:0006118; P:electron transport; IEA.					
DR	InterPro: IPR001128; Cytochrome P450.					
DR	InterPro: IPR008072; EP450_Cyp3A.					
DR	Pfam: PF00067; P450; 1					
DR	PRINTS; PR00687; EP450IICYP3A.					
DR	PRINTS; PR00385; P450.					
DR	PROSITE; PS00086; CYTOCHROME_P450; 1.					
DR	Heme; Monooxygenase; Oxidoreductase.					
SW	SCQUENCE 503 AA; 57085 MW; 5A17128CA9E50914 CRC64;					

Query Match 14.1%; Score 403.5; DB 6; Length 503;

Query Match	14.18;	Score 403.3;	22.0
Best Local Similarity	27.88;	Pred. No. 4.4e-21;	

Matches 152; Conservative 81; Mismatches 191

[illegible]

12 GLAAFSWASIA--FFSLYLAPRRS-----SLYNLQGNH

[illegible]

6 GFSTETWLLATSLVLLYLYGTYSHGLFKKLGIPGPRP



```

QY 64 AKYREKYSTLRFGAGI--AGAPVLNSTDPKVFNVHM-KEAY-----DYPKFGMAARVL 113
DB 62 KKCFOQYG---KMGVYDGRQPLLAIVTDPNMIKSVLKCYVSFTNRRSFGPLGAMENAL 118
QY 114 RATGQGVVTAAGEAHKRRRIMIPSLSAQAVKSNVPIFEKGMELVDKMWDAEAKOMA 173
DB 119 SL-----ADEBWKIRILLSTFTSGKLEKMFIIISHYGOLLVSNURKE-AEKG-- 167
QY 174 VGESAGEKATRLTEGVDVQWVGRATLDVMAAGFDYKSDSLQNK-----TNELYVAF 228
DB 168 -----KPYTKMDIFGAYSMVITSTAFGWNIDSLNNPQDPFVENSCKLLK 212
QY 229 VGLTGAFAPTLDSFKAIMWDF-----VYPF-RTMKRRHEIPLTQGLAVSRVVG 275
DB 213 FSEFPFLSLIFFPFLPIFEVLNITLFPKSSVNFFTKSVKMKESRLT----- 262
QY 276 IELMEQKQAVLGSASDAQVDKQVGRDILSLVRAANIAANLPSQKLSDEEVLQISN 335
DB 263 -----DOQKERV-----DLQLMINSONSKEMDPHKSLSNEELVGPGLI 301
QY 336 LRFAGYETSTVLTWTFHRLSEDKAVQDKLREIEICQI--DTDMPTLDELNALPYLEAFVK 393
DB 302 FIFAGYETTSALSLLAYELATHPDVQKQLOEIEATEPKNKAPPTYDALAQMEYLDWVN 361
QY 394 ESLRLDPPSPYANRECLKDEDEPILEAFVIGRDSGVINEVRITKGTWMLPLFNINRSKF 453
DB 362 ETURLYPIARLERACKDVE-----IHGVFVPGTGVVVPFVFLHRDPD 406
QY 454 IYGEDAEFRPERWLEDVTDLSNLSIEAPYGHQASFISSPRACFGWRFAVAEMKAFLEVT 513
DB 407 LWPE-PEEFPERFSKKNKDTIN---PVTY-LPFGTGPRNCIGWRFAVMMKLALVKVL 460
QY 514 RVQFEP 520
DB 461 QNPSFKP 467

RESULT 7
Q9LUD2 PRELIMINARY; PRT; 515 AA.
AC Q9LUD2;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Cytochrome p450 (AT3g14620/MIE1.12).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Nakamura Y.;
RX MEDLINE=20277480; PubMed=10819329;
RN [3]
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carinci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamaya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,

```

```

RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA "Arabidopsis cDNA clones."
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB023038; BAB02394.1; .
DR EMBL; AY052208; AAK97679.1; .
DR GO; GO:0004497; P:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 515 AA; 5843 MW; 3F45FCF7686F55A1 CRC64;

Query Match 14.1%; Score 403; DB 10; Length 515;
Best Local Similarity 26.5%; Pred. No. 5e-21;
Matches 134; Conservative 87; Mismatches 191; Indels 94; Gaps 19;

QY 42 NHTNYFTGNFLDILSARTGGEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVNKEAY 101
DB 76 NLTDYTHRVNPLIQ-QIVKHGK-----TSYWMGPIASVIV---TKBEHIKDLNRYV 126
QY 102 DYPKFGMAARVLRATGCGVVTAGGEAHKRRRIMIPSLSAQAVKSNVPIFEKGMELVD 161
DB 127 DFPKPPVHPIVELEAT--GVALYEKWSKRRKIINPSFHLKLIKIMIPAFYESCSEMIS 184
QY 162 KMEDIAAKOMAVGESAGEKATLETEGVYDKVQVGRATLDVMAAGF--DYKSD---- 215
DB 185 KW-----EKLVTGEGSNE-----IDVPIGLDITSDVISRTAFGSSVEBGRIF 229
QY 216 SLQNKTELYVAFVGLTDFGAPTLDSFKAIMWDFVYPFRTMKRHEIPLTQGLAVSRVVG 275
DB 230 ELQEEQGRVL-----KALELAFIPGMRFPLTKNNLRMQ-----INKEVK 270
QY 276 IELME--OKQAVLGSASDAQVDKQVGRDILSLVRAANIAANLPSQKLSDEEVLQAI 333
DB 271 SRLREIMKQ-----RGMDTGEAPKNDLIGLLESNG-----DHGMSIEDVVEEC 317
QY 334 SNLLFAGYETSTVLTWTFHRLSEDKAVQDKLREIEICQI--DTDMPTLDELNALPYLEAF 391
DB 318 LRFAGAQETTAVLLVMTIMLSHHQKQVQDQAREIILKGNKNFNFALSLSKTMSMI 377
QY 392 VKESRLDPPSPYANRECLKDEDEPILEAFVIGRDSGVINEVRITKGTWMLPLFNINRS 451
DB 378 LNEVLRLYPPGILLGRVTEK-----ETKLGED-----MTLPGGAQVVIPLVMVHRD 423
QY 452 KFIYGEDAEFRPERWLEDVTDLSNLSIEAPYGHQASFISS--GPRACFGWRFAVAEMKAF 508
DB 424 PELMGEDVHEFNPERFADGISKATK-----NQVSFLPFGWGRPCPGQNFALMEAKWA 476
QY 509 LEVTLRRVQFE--PIISHPEYEHITL 532
DB 477 LVLLIQRFSELSPSYTHAPHTVLT 502

RESULT 8
Q8WIE1 PRELIMINARY; PRT; 515 AA.
AC Q8WIE1;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE AT3g14620/MIE1.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Banh J., Bowser L.,

```

RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Iehida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,  
 RT "Arabidopsis ORF clones";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR ENBL; AF443860.1; -  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; Oxidoreductase.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 515 AA; 58616 MW; 9F4AD6C768648CA1 CRC64;

Query Match 14.1%; Score 401; DB 10; Length 515;  
 Best Local Similarity 26.5%; Pred. No. 7e-21;  
 Matches 134; Conservative 86; Mismatches 192; Indels 94; Gaps 19;

OY 42 NHTNYFTGNFLDILSARTGEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNNHYMKEAY 101  
 DB 76 NLTDYTHRVPLIQ-QTVKDHCK-----TSYMMGPIASVIV--TKPEHIKDLNVRVY 126  
 OY 102 DYPKPGMAARVLRIATGQGVVTAEGEAHKKHRRIMIPSLSAQAVKSMVPIFEKGMELVD 161  
 DB 127 DFPKPPVHPVIVELFAT--GVALYEGEKWKKHRIINPSFHLKXIKIMIPAFYSCSEMI 184  
 OY 162 KMEDEAAEKDMVAGESAGEKKAIRLETETGVVDKDWGRATLDVMALAGF--DYKSD---- 215  
 DB 185 KW-----EKLVTGGSNE-----IDWPFYGLDILTSVISTAFGSSVEEGKRF 229  
 OY 216 SLQKNTNELYVAVGLDGPAPLIDSEKALIMWDFVYFTMKRRHEIPTQGLAVSRVVG 275  
 DB 230 ELQEEQRRVL-----KALELAFIPGRFLPTKSNLMRQ---INKEVK 270  
 OY 276 IELME--QKQAVLGSSDAVDKQVQGRDILSLVLRANIANLPESQKLSDEVLQAI 333  
 DB 271 SRLREITMKQ-----RGMTGEAPNDLLGILLESNG-----DGMKIEDVVEEC 317  
 OY 334 SNLLFAGYETSSVLTWFMHRLSEDKAVQDKLEETICQI--DTDMPTLDNALPYLAEAF 391  
 DB 318 RLPHFAGETTAVLLWTMTLSHHQKQDQAREILKVIKGNKPNFALSRLKXTMSMI 377  
 OY 392 VKESIRLDPPSPYANRECLKDEPIPLAEPVIGRDSVINEVRIKGTWMLPLFNINRS 451  
 DB 378 LNEVRLYPPGILLGRVEX-----ETKGED-----MTLPGGAQVIVPLVMVRHD 423  
 OY 452 KFIYGEDAEEPRRWLEVDYDLSINSIEAPYHQASFIS---GPRACFGWRFAVEMKAF 508  
 DB 424 PELWGEDVHEFNPERFADGISKATK-----NQVSLFPGWGRFCPCGNFALMEAKWA 476  
 OY 509 LFTVLRVQFP--PIISHPEYEHITL 532  
 DB 477 LVLLIQRFSELSPESTHAPHTVLT 502

RESULT 9  
 QASR3  
 ID QASR3 PRELIMINARY; PR7: 519 AA.  
 AC QASR3; OS0728;  
 DT 01-JUN-2001 (TREMELrel. 17, Created)  
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE At2g45960/F14M4.21 (Putative cytochrome P450).  
 GN At2g45960.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,  
 RA Bower L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,  
 RA Iehida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Davis R.W., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,  
 RA "Arabidopsis cDNA clones";  
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,  
 RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;  
 RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA Lin X.;  
 RA STRAIN=cv. Columbia;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN=cv. Columbia;  
 RA Town C.D., Kaul S.;  
 RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,  
 RA Bower L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Iehida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RA "Arabidopsis ORF clones";  
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RL [6]  
 RN -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR ENBL; AF367329; AAC32916.1; -  
 DR ENBL; AC004411; AAC34227.2; -  
 DR ENBL; AY091688; AAW10287.1; -  
 DR PIR; T02191; T02191.  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; Cytochrome\_P450; 1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 519 AA; 59603 MW; 3B053C83F917457A CRC64;

Query Match 14.0%; Score 400.5; DB 10; Length 519;  
 Best Local Similarity 23.5%; Pred. No. 7.7e-21;  
 Matches 138; Conservative 110; Mismatches 231; Indels 109; Gaps 20;

OY 1 MFILVLTGALGAAPSWASIAFFSLYLAIPRSSLYNLOGPHNYFTGNP----- 51  
 DB 9 LIVLILI---IGLRFKAFMILVWHFVLTRELKNGQISGPNY-RIFYGNSIEIKMKRE 64  
 OY 52 --LDILSARTGE-----EHAKYREKYSTLRFAGIAGAPVLNSTDPKVFNNHYMKEAYD 102  
 DB 65 SHLSILDPSSNDIFFPRILPHYQKWSQYGETLYWN-GTEPRICISDELAKTMSLNKLG 123  
 OY 103 -YKPGMAARVLRIATGQGVVTAEGEAHKKHRRIMIPSLSAQAVKSMVPIFEKGMELVD 161  
 DB 124 FFVKSKARPEAVKLVGSKGLVFIEGADWVRHRIINPAFSDIRLKIMTVAVDCTL---- 179  
 OY 162 KMEDEAAEKDMVAGESAGEKKAIRLETETGVVDKDWVG---RATLDVMALAGFDYKSDSL 217  
 DB 180 KMLEE-----WRKSTKEETHPKIKENEEFQRLTIATISAFG----- 221

```

QY 218 ONKTNELVAVFGLTGDPAFTDSEKAIMMDFVFPYFMTKRRHEIPLTQGLAVSRVVGIE 277
DB 222 -----SSYVEGIEVFSQM-ELKRCYTTSLNQSVIPGTQYLPPTPSNIRYM 265
QY 278 LMEQKQAVLGASDAQVKKDVOGRDITLSLVANIAANLPSQKLSDEVLAIQISNLL 337
DB 266 KLERKMSINSIKRITSRLQSKSDVGDLLGLLXAYNTEG--KERRKSIIEIHCRTFF 323
QY 338 FAGYETSIVTWMFHRLESDKAVQDKLREI---CQIDTMDPTLDELNALPYLEAFVKE 394
DB 324 FGHETTINLAWITMLLSLHQQWQEKLEIIPKECGKE-KTDSSETFSKLKMNVMIME 382
QY 395 SURLDPPSPYANRECLDEDEPILEAPVIGDGSV---INEVRITKGTWMLPLFNINRS 451
DB 383 SURLYGP-----VSALAREASVNIKGLDEIPKTTVVIPLLKQHS 424
QY 452 KFIYGEDAEFRPERLWEDVTDLSNIEAPYGHQASFSIGPRACFGWRFVAEMKAPLFV 511
DB 425 KTLWGSADAKFNPMFANGVSRANHPNA----LLAFSGVPRACIGQNFVMEIAKTVLT 480
QY 512 TLRRVQFPPIISHPIYEH--TLLISRPRIVGREKEGYQMLQVKE 557
DB 481 ILQRFPRFISLQD--EYKHTPDVNTIQD-----YGLFVMLQPLE 518

RESULT 10
Q8LHV0 PRELIMINARY; PRT; 519 AA.
AC Q8LHV0;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN P0035D09.13.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone:P0025D09."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF004264; BAC10362.1; -.
DR Gramene; Q8LHV0; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 519 AA; 58547 MW; 77E9E9C1428FE38 CRC64;

Query Match 13.9%; Score 397.5; DB 10; Length 519;
Best Local Similarity 26.0%; Pred. No. 1.3e-20;
Matches 155; Conservative 80; Mismatches 225; Indels 137; Gaps 22;

QY 10 ALGLAIF-----SWASIAFSLYLAP-----RSSLNLOGPNHTNYFTGNLDI 54
DB 2 AMGLLAWMAAAAAAASLASWAFSAVHLVWREPAISRRLRAQGVGGY-RFFSGNLGEI 60
QY 55 LSAR-----TGEHAKYREKYGSLTRFAGIAGAPVLNSTDPKVFNVH 96
DB 61 KRFRGAGVNLVNSHDFLPVQPHFRKWIPLYGRTFLY-WFGAQPNICLADYVMWQV 119
QY 97 MKBAYD-YPKPGMAARVLRIATGDGVVTAEGSAHGRHRIMIPSLSAQAVKSWVPIFEK 155

Db 120 LSRDTGIYKXNLNPHFVRL-LGKGLVLTDGDEWKEHRRKVVHPAFNMCKLKWMT----- 172
QY 156 GMELVDQMDEDAEAKDNVAGESAGEKATRELETGVDVKDWVGRATLDIMALAGF--DYK 213
DB 173 -----MTNSDCRSRMWSEWELAAKGLVE---IELSRREELTADVISHTAFGSSYK 223
QY 214 SDS---LQNTNELYVAFVGLT---DQFA--PTLDSFKAIMMDFVFPYFTM-----KRKH 260
DB 224 EGKQVFLAQRELQFLAFSTFLTVQIFGFSVLPIMKFKT--WSLDKKVRGMLMDIKTRH 281
QY 261 EIPITQGLAVSRVVGIELMEQKQAVLGASDAQVKKDQV--QGRDILSLVANIAML 318
DB 282 A-----NKDVAGYGNLGLGLEACAPHG 306
QY 319 PESOKLDEEVLAIQISNLLFAGYETSIVTWMFHRLESDKAVQDKLREIEIC-QIDTMDP 377
DB 307 ESCQLSDEHDIIECKTFFPAGHDTTSHLLTWMLFLLSTHDPWQEKLEIEIAECGDKVP 366
QY 378 TLDELNALPYLEAFVKESSLRDLDPSPYANRECLKQSDFIPLAEFVIGRDGVSINEVRIYK 437
DB 367 TGDMLNKLKVMNMFLETLRLYSPVSLIRRKVDTDIE-----LGGIKMPE 411
QY 438 GTWMLPLFNINRSKFIYGEDAEFRPERLWEDVTDLSNIEAPYGHQASFSIGPRACFG 497
DB 412 GALLTIPIATHIRDXKEVWGDEDADEFPERFENGVTAAKHNA-----LLSFSSGPRSCIG 467
QY 498 WRFVAEMKAPLFVTLRRVQF--EPIISHPEYEHITLIISRPRIVGREKEGYQMLQ 552
DB 468 QNFAMIEAKAVIAMILQRFSTLSPKYVHAPTVDITL-----RPKYGLPMILK 515

RESULT 11
Q86SK3 PRELIMINARY; PRT; 503 AA.
AC Q86SK3;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Cytochrome P450.
GN CYP3A43/CYP3A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21839017; PubMed=11726664;
RA Finta C., Zaphiropoulos P.G.;
RT "Intergenic mRNA molecules resulting from trans-splicing."
RL J. Biol. Chem. 277:5882-5890(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Zaphiropoulos P.G.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR GO; GO:0006712; P:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PROSITE; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 503 AA; 57394 MW; 4141F95B28B34A5E CRC64;

Query Match 13.8%; Score 394; DB 4; Length 503;
Best Local Similarity 28.7%; Pred. No. 2.2e-20;
Matches 142; Conservative 83; Mismatches 177; Indels 92; Gaps 20;

QY 47 FTGNFLDILSARTG-----EEHAKYREKYGSLTRFAGIAGAPVLNSTDPKVFNVH-XE 99
```

```

Db 46 FLGN---ILSVHKGFCMFEDMECHKYGVKMGF---YDG--QOPVIAITDDPMIKTVLKE 97
QY 100 AYDY---PKPGMAARVIATGDGVVTAAGBAHKEHRRIMIPSLSAQAVKSWPIFLK 156
Db 98 CYSVFTNRFPFGVGMKSA---ISIADEEWKRLSLLSPTFTSGKLKEMVPIIAQYG 153
QY 157 MELVDKMDAEDAEKMAVGESAGEKATRLT-EGVDVQKDWVGRATLDWALAGFDYKSD 215
Db 154 DVLVRNLREA-----ETGKPVTLKDVFGAYSMDVITSTSGFNID 194
QY 216 SLQNKTNELVAVFVGLTDFAPLDSFKAIM-WDFV-PYFRTMK-RRHEIPTLOGLAV-- 270
Db 195 SLNPNPD-----PFVENTKLLRFDFLDPFPLSITVFPFLIPILVNICV 240
QY 271 -SRVGIEMEKQKQAVLGSQDQAVDKVQVGRDILSLVYRANIAANLPSQKLSDEEV 329
Db 241 FPREVTNFRKSVKR-----MKESLDEPTQKRVDFLQIMDSQNSKETESHKALSLEL 295
QY 330 LAQISNLLFAGYETSTVLTWTFHRLSEDAVQDKLREEICQI--DTDMPTLDELNALPY 387
Db 296 VAQSIIFIFAGYETSSVLSFMYELATHPDVQOKLQEIIDAVLPNKAPPTDYDVLQMEY 355
QY 388 LEAFVKESLRLDPPSPYANRECLKQEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFN 447
Db 356 LDMVNETLRLPFIAMRLERVCKDVE-----INGMFIKGVVWVIPSYA 400
QY 448 INRSKIYGEDAEERPRERWLEDVTDLSNS-IEAPYHQASFIQSPRACFGWRFAVEMK 506
Db 401 LHDP-KYVTEPEKFLPERFSKKNKDNDPIYTFEG-----SGPRNCIGMRPALNMNK 453
QY 507 AFLFVTLRRVQEP 520
Db 454 LALIRVLQNFSEKP 467

RESULT 12
Q9FEEL PRELIMINARY; PRT; 544 AA.
AC Q9FEEL;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN P0688A04.9 OR P0006C01.24.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0688A04.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0006C01.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AP002839; BAB19103.1;
DR EMBL; AP002744; BAB19082.1;
DR Gramene; Q9FEEL;
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.

```

```

KW Heme; Monooxygenase; Oxidoreductase
SQ SEQUENCE 544 AA; 60867 MW; 273EAF5968D1A024 CRC64;

Query Match 13.8%; Score 393.5; DB 10; Length 544;
Best Local Similarity 26.4%; Pred. No. 2.7e-20;
Matches 159; Conservative 83; Mismatches 215; Indels 145; Gaps 25;

QY 6 LLTGALGLAAPSASIAFFSLYLAPR-SSLYNLQGNHNY-----FTGNFLDLSAR 58
Db 31 LLGGVAALLLVWAAQMLEWAWLAPRMERALRAQGLRGTOYRFLFGDLDLRLVTAAR 90
QY 59 TG-----EEHAKYREKYGSLRFPAGIAGAPVNLNSTDPKVFNVHV 96
Db 91 SKPVMDRPHDFIPRVAPLLHRALEHGR-----VSFTFGPMRPTVITDPLVREV 142
QY 97 MKEAY-DYFKGMAARVLRATGCVVTAAGBAHKEHRRIMIPSLSAQAVKSWPIFLK 155
Db 143 LSNFGHEFKTKATRLSKLLVG-GLVILHGEKVVHRRIMNPFHAEKLRMLPAPSAS 201
QY 156 GWELVDKMDAEDAEKMAVGESAGEKATRLTETEGVDVQKDWVGRATL--DWMALAGFDYK 213
Db 202 CSELIGRW-----ENAVAASVGKAE-----LDI--WPDFQNLSDGVISRAAFGVR 244
QY 214 SDS-----LQNKTNELVAVFVGLTDFAPLDSFKAIMWDFVYF-----RTMKRHEI 262
Db 245 RHGGRQIFLQAEQAEHLV-----QSFRS---NYIPGLSYENNRNKK-----283
QY 263 PLTQGLAVSRV-----GIELMEQKQAVLGSQDQAVDKVQVGRDILSLVYRANIAAN 317
Db 284 -----AIDRIKILRGI-IEKQKATKNGAS-----KD---DLLGLLQSNMDYY 326
QY 318 LPESOKLSD-----BEVLAQISNLLFAGYETSTVLTWTFHRLSEDAVQDKLREEICQI- 372
Db 327 SDEGKSKGVTVEEIIDECKLFYFAGMETTAVLLTWTWVALSNHPEWQDRAREEILQVF 386
QY 373 DTDMPFLDELNALPYLEAFVKESLRLDPPSPYANRECLKQEDFIPLAEPVIGRDSVINE 432
Db 387 GRNKPDINGVSRKLVTVMLHEVLRLYPFVVMNRRTYKIE-----LGG 431
QY 433 VRITKGTWMLPLFNINRSKIYGEDAEERPRERWLEDVTDLSNSIEAPYHQASFIQSP 492
Db 432 VRYPAGVMLSPLVLFIRHDAAGHGDAGEFDPGRFAEGVARACKDPCA--GAFPPFSWGP 489
QY 493 RACFGWREAVAEKMAFLPVTLLRRVQEB--PIISPEYEHITLLISRPRIVGREKEGYQMR 550
Db 490 RICIQNFALLEAKVALGNILQRFAPLSPAYAEAPYVITL-----HPQHGVPVR 540
QY 551 LQ 552
Db 541 LR 542

RESULT 13
Q8CJF2 PRELIMINARY; PRT; 497 AA.
AC Q8CJF2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome P450 3A.
GN CYP3A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Matsubara T., Nagata K., Yamazoe Y.;
RT "Isolation and characterization of a novel rat CYP3A form.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB084894; BAC23085.1;
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d...; IEA.

```

```
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 497 AA; 57009 MW; 8C034C25C476F452 CRC64;

Query Match      13.7%; Score 392; DB 11; Length 497;
Best Local Similarity 26.4%; Pred. No. 3e-20;
Matches 140; Conservative 81; Mismatches 176; Indels 134; Gaps 20;

QY 35 LYNLOGPHTNY-----PTGNFIDLSARTG-----BEHAKYREKYGSTLRPAG 78
Db 22 LLYGTSTHGNFKLIGSGPKLPFVGN---ILAYRHGFWEFRHC---HKYGDINGFYE 76
QY 79 TAGAVLNSTDPKVFNVHM-KEAY-----DYPKPGMAARVLRIRATGCGVWTAEGEAHK 130
Db 77 -GRQILAITDPDIKTIVLKECYSTFNRRSFGPAGILKKAITL-----SEDEWK 127
QY 131 RHRRIMPSLQAQAVKSMVPIFEKGMELVDKMDAEDAAEKDMAGVESAGEKKAIRLETEG 190
Db 128 RLRTLLSTFTSGKLKMFPI-INQYADLLVKNVKEAEKG-----NP 169
QY 191 VDVKDWGVRATLDVWALAGFDYKSDSLQNTN-----ELYVAFVGLTDGF 235
Db 170 ITMKDIFRAYSMDVITGTFSGVNDLNNPQNPVQVKVKKLLKFNFLDPFLSVLPFL 229
QY 236 APTLDSFKAIMW--DFVPYFTMKRRHEIPLTQGLAVSRVRVGIEMQKQAVLGASDQ 293
Db 230 TVPFAFDITVPFKDVKMFRT-----SVERMKENR----- 260
QY 294 AVDKDVQGRDILSLVRANLANLPSQKLSDEVLQISNLIFAGYETSTVLTWMPH 353
Db 261 -MOEKVKQBLDFQLMINSQSGGKESHQGTUVEIQAQSFIFPAGYETSSALSPLY 319
QY 354 RLSEDKAVQDKLREICQIDTMP-----TLDELNALPYLFAFKESLRLDPPSPYANRE 408
Db 320 LLATHPDQLKQKQDE--IDAALPNKAPVTVYDVLVEMBYLDMVNLNLTFLFPVGGRLERV 376
QY 409 CLKDSDFTPLAEFVIGRDSVINEVRIKGTMMVLPFLNINRSKFIYGEDAEFRPERWL 468
Db 377 CKKQVE-----INGVIFPKTVWVPTFALHKDKPCWPE-PEEPCPERFR 420
QY 469 EDVTSLSNS-IEAPYCHOASFTSGPRACFGWRFAVEMKAFVLTLRVQF 518
Db 421 KKNQDSINPYIYLPQ-----NGRNCIGMRFAIMNNKIALVRVLQNFSF 465

RESULT 14
Q8ISJ7 PRELIMINARY; PRT; 492 AA.
AC Q8ISJ7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome P450 CYP4L4.
GN CYP4L4.
OS Mamestra brassicae.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Hadeninae; Mamestra.
OX NCBI_TaxID=55057;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Antenna.
RX MEDLINE=21956093; PubMed=1200647;
RA Maibeche-Coise M., Jacquin-Joly E., Francois M.C.,
RA Nagnan-Le Meillour P.;
RT "cDNA cloning of biotransformation enzymes belonging to the cytochrome
RT P450 family in the antennae of the noctuid moth Mamestra brassicae.";
RL Insect Mol. Biol. 11:273-281(2002).
```

```
DR EMEL; AY063501; AAL48300.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 492 AA; 56219 MW; 4F69897198D9D01B CRC64;

Query Match      13.6%; Score 389; DB 5; Length 492;
Best Local Similarity 22.9%; Pred. No. 5e-20;
Matches 139; Conservative 105; Mismatches 182; Indels 180; Gaps 24;

QY 1 MFILVLTGALGLAAFSWASIAFFSLYLAPRSSLYNQPN-----HTNYF---TGNF 51
Db 6 VFVVVLA-----LLVSWISLV-----RVSRRENVPGPSPLPLVGNALHFLVVKSSF 52
QY 52 LDILSATGEHAKYREKYGSTLRFAGIA-----GAPVLNSTDPKVFNVHMKEA 100
Db 53 LNLVQ-----RLSEKYGHAFRVHFFSTFYVVICHSKYAEPLVSSTB-----HITK-- 97
QY 101 YDYPKPGMAARVLRIRATGCGVWTAEGEAHRRHRRIMPSLSAQAVKSMVPIFEKGMELV 160
Db 98 -----GRSFSPLTCMLGQGLLTATQGRWKSHERKELTPAFHFNILQNFLEPVCCKQRILT 151
QY 161 DKM--MEDAAEKDMAGVESAGEKKAIRLETEGVVDKDWGVRATLDVWALAGFDYKDSL- 217
Db 152 EKIRGMADGRPIDM-----FPIIALAALDNVNESIM 182
QY 218 -----QNKINELVAFVGLTDGFAPLDSFKAIMMDFVPYFTMKRRHEIPLTQGLAVS 271
Db 193 GVCMDAKHQSSEV-----KSIELSIV-----TM--RMQIPFFGEDAIF 222
QY 272 RRVGIELMEQKQAVLGASDQAV-----DKDVQGRD-----ILSLVRA 312
Db 223 NLLPKTKQDKALKVLHGQTKNVINARRAELKKNITLNDSSDICTKKNKHTFLDLLLA 282
QY 313 NIANLPSQKLSDEVLQISNLIFAGYETSTVLTWMPHRLSEDKAVQDKLREI--- 369
Db 283 EI-----DGKIDDESVRVEDTFWFEGHDTTISGIVYTLHCLSKRQDVQEKIYELKTI 337
QY 370 --CQIDTMPITDELNALPYLFAFKESLRLDPPSPYANRECLKDEDFIPLAEFVIGRQ 427
Db 338 FGSEIHRD--PYHELOQMVKYLELVIKESMLRFPFVPLIERRIMKDCE----- 383
QY 428 SVINEVRIKGTMMVLPFLNINRSKFIYGEDAEFRPERWLEDVTSLSNSIEAPYCHOAS 487
Db 384 --VGSGLVKGTSTVYVNIFOIQOPDLF--DDPLEFRPERF-----EAPLKNPFS 429
QY 488 ---FTSGPRACFGWRFAVEMKAFVLTLRVQFEPFIISHPEYEHITLIISRPRIVGREK 544
Db 430 WLAFSAGPRNCIGQKFAWMLKITISIVKNFFILPAABEPFELS-----ADLVLRSK 481
QY 545 EGYQMR 550
Db 482 NGVNVK 487

RESULT 15
Q8LQ77 PRELIMINARY; PRT; 523 AA.
AC Q8LQ77;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN B1051E10.19.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN=cv. Nipponbare;  
RA Sasaki T.; Matsumoto T.; Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
RT clone: B1051E10."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AP003764; BAB93411.1; .  
DR Gramene; C81Q77; .  
DR GO; GO:0004497; F:monooxygenase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; p450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME P450; 1.  
DR Heme; Monooxygenase; Oxidoreductase.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 523 AA; 58192 MW; DSB6C6C0F0965798 CRC64;

Query Match 13.6%; Score 387; DB 10; Length 523;  
Best Local Similarity 24.8%; Pred. No. 7.7e-20;  
Matches 143; Conservative 94; Mismatches 245; Indels 94; Gaps 19;

Qy 3 ILVLTGAL---GLAAPSWSIAFFSLYLAPRESSLYNQGNHNYFTGNELDILSART 59  
Db 10 ILVVLVSRLLTSALVHLWKPAYATKLRG-----QGITGPKY-RLFVGSLSPEIKRMKA 62  
Qy 60 G-----EEHAKYREKYGST-LRFAGIAGAPVLNSTDPKVFNNHYMK 98  
Db 63 AAAADEVAAAGAHSHDFIPIVLPOHSHKWATDHGKTFLYWLGAVPAVSLGRVE-QVKQVLE 121  
Qy 99 BAYDYPKPGMAARVLRTATGCVVTAEGEAHKEHRRIMIPSLSAQAQVKSVMVFIFLEKME 158  
Db 122 RTGSFTKNYMNAN-LEALLGKGLIANGEDWERHKKVHPAFNHDKLFKMSVYMAESVES 180  
Qy 159 LVDXWMEDEAEKMAVGESAGEKATRETEGVVDVVDWVGRATLDVWALAGFDYKSDSIQ 218  
Db 181 MVQRWQSQI-----QQAGNQVE-----LDLSRELSELTSVITRSAP-----GSSH 222  
Qy 219 NKTNELYVAVGLTD-GFAPTLDSFKAIMWDFVYPRTWKRRHEIPLTOGLAVSRVGLIE 277  
Db 223 BEGKEVYQAQKELQELAFSSLDVPAVFLRKLPIIRGNTRAHQ-----VKKSRITLME 277  
Qy 278 LMEOKKQAVLGSASDAQVDKQVGRDILSLVYRANIAANLPESOKLSDEEVLQISNLL 337  
Db 278 IIEGRLAKEAABAG-----YGSLLGLMEARALEREENGLVLTTOEIIDCKTFF 329  
Qy 338 FAGYETSTVLTWMFHLSEDKAVQDKLREIICQIDTD-NPTLDELNALPYLEAFVKESL 396  
Db 330 FAGQDTTSHLVWTFMFLSSNAQWQDKLREEVLTVCQDAIPTDPMANRLKLVNMVILESL 389  
Qy 397 RLDPSPYANRECLKDEDFFLAEPVIGRDSVINEVRIKGTWVWMLPLFNINRSKFIYG 456  
Db 390 RLYSPVVIIRIAGSDID-----LGNLKIPIKGVLSIPIAKIHRDRDWG 434  
Qy 457 EDAAEFPRERWLEDVTDLSNLSIEAPYGHQAFSGPRACFGMRFAVAEMKAFVLTLRV 516  
Db 435 PDADFNFPARFKNGVRAASYPNA-----LLSFSQGPREGCIGQTFAMLESQIAIAMILQRF 490  
Qy 517 QF--BPIISHPEYEHITLISRP-----IVGREKEG 546  
Db 491 EFRUSPSYVHAPMEAITL---RPRFGLFVVLRLNQ 523



# STIC SEARCH RESULT FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 04:33:00 ; Search time 170 Seconds  
(without alignments)  
6306.854 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932

Sequence: 1 gaattcgacagagccacc.....agcggctgtgcgaattc 1932

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1932	100.0	1932	4	US-09-518-386B-2
2	353.2	18.3	3969	4	US-09-518-386B-4
3	48.4	2.5	1512	4	US-08-277-031B-4
4	48.4	2.5	2059	4	US-09-023-655-1062
5	48.4	2.5	2759	4	US-09-144-367-1
6	47.4	2.5	2160	4	US-09-716-129-12
7	46.8	2.4	1712	4	US-09-148-545-106
8	46.8	2.4	1822	4	US-09-148-545-105
9	45	2.3	1831	3	US-08-948-564-15
10	44	2.3	1831	4	US-09-221-013A-7
11	43.6	2.3	352	4	US-09-144-367-8
12	42.8	2.2	1333	4	US-09-372-422A-9
13	42.8	2.2	2407	3	US-09-370-807-7
14	42.8	2.2	2407	4	US-09-921-259-7
15	42.6	2.2	246240	2	US-08-724-394A-21
16	42.6	2.2	246240	2	US-08-724-394A-20
17	42.6	2.2	246240	2	US-08-724-394A-22
18	42.4	2.2	396	4	US-09-640-173-19
19	42.4	2.2	396	4	US-09-713-550-19
20	42.2	2.2	47	2	US-08-778-494B-114
21	42.2	2.2	323	4	US-09-621-976-10374
22	41.6	2.2	2202	4	US-09-396-149-3
23	41.4	2.1	1126	4	US-09-389-956-7
24	41.4	2.1	1507	4	US-09-453-323-1
25	41.4	2.1	7218	1	US-08-232-463-14
26	41.2	2.1	1024	4	US-09-328-475C-107
27	40.8	2.1	44848	4	US-09-435-739-42

28	40.6	2.1	1883	4	US-09-419-679-13	Sequence 13, Appli
29	40.2	2.1	947	4	US-09-573-395A-8	Sequence 8, Appli
30	40.2	2.1	5152	4	US-10-204-708-48	Sequence 48, Appli
c	40	2.1	155	4	US-09-821-976-11009	Sequence 11009, A
31	40	2.1	1260	4	US-09-674-741-5	Sequence 5, Appli
32	40	2.1	2581	2	US-09-013-634-1	Sequence 1, Appli
33	40	2.1	2745	4	US-09-833-381-1410	Sequence 1410, Ap
34	40	2.1	2745	4	US-09-023-655-1405	Sequence 1405, Ap
35	39.8	2.1	1599	4	US-09-800-729-56	Sequence 56, Appli
36	39.8	2.1	2099	4	US-09-621-976-9843	Sequence 9843, Ap
37	39.6	2.0	214	4	US-09-640-173-16	Sequence 16, Appli
c	39.6	2.0	396	4	US-09-713-550-16	Sequence 16, Appli
c	39.6	2.0	1001	4	US-09-671-317-84	Sequence 84, Appli
40	39.6	2.0	2114	4	US-09-130-491-7	Sequence 7, Appli
41	39.6	2.0	3289	4	US-09-904-615-11	Sequence 11, Appli
42	39.6	2.0	6200	4	US-09-439-923-1	Sequence 1, Appli
43	39.6	2.0	6200	4	US-09-711-202A-1	Sequence 1, Appli
44	39.6	2.0	6200	4	US-09-711-202A-1	Sequence 1, Appli
45	39.6	2.0	6200	4	US-09-711-205A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-518-386B-2  
; Sequence 2, Application US/09518386B  
; Patent No. 6365386

; GENERAL INFORMATION:

; APPLICANT: HOSHINO, Tatsuo  
; APPLICANT: OTIMA, Kazuyuki  
; APPLICANT: SETOGUCHI, Yutaka  
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
; CURRENT APPLICATION NUMBER: US/09/518.386B  
; CURRENT FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: EP 99104668.1  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EP 00101666.6  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1932  
; TYPE: DNA  
; ORGANISM: Phaffia rhodozyma  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (33)..(1706)  
; NAME/KEY: polyA\_site  
; LOCATION: (1871)  
; NAME/KEY: rRNA  
; LOCATION: (14)..(1891)  
US-09-518-386B-2

Query Match 100.0%; Score 1932; DB 4; Length 1932;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAATTCGGCAGAGGCCACCTACTTCTCCATATGTTTCATCTTGGTCTTGTCTCACAGGTG	60
DB	1	GAATTCGGCAGAGGCCACCTACTTCTCCATATGTTTCATCTTGGTCTTGTCTCACAGGTG	60
QY	61	CTTTAGGCTGGCTGCTTTCTCATGGGCATCCATAGCGTTCCTTCAGTCTTTTACCTCGCTC	120
DB	61	CTTTAGGCTGGCTGCTTTCTCATGGGCATCCATAGCGTTCCTTCAGTCTTTTACCTCGCTC	120
QY	121	CGAGCGCATCTTCACTGTATTAACCTTCAGGCCCGCATCATACCACTACTTTACAGGCA	180
DB	121	CGAGCGCATCTTCACTGTATTAACCTTCAGGCCCGCATCATACCACTACTTTACAGGCA	180
QY	181	ATTTTITAGACATCTCTCAGCTGTAGAGTGAAGAGCATGCGAAGTACAGAGAAAAT	240
DB	181	ATTTTITAGACATCTCTCAGCTGTAGAGTGAAGAGCATGCGAAGTACAGAGAAAAT	240



241 ACAGAGACACCTCGGTTTCTCGGATCGCTGGAGCACCGCTTGAATCTCGACCGATC 300  
Db  
241 ACAGAGACACCTCGGTTTCTCGGATCGCTGGAGCACCGCTTGAATCTCGACCGATC 300  
Qy  
301 CGAAAGTCTTCAACCATGTGATGAAGAAGCTACGACTATCGAAACCTGTGTATGCGCG 360  
Db  
301 CGAAAGTCTTCAACCATGTGATGAAGAAGCTACGACTATCGAAACCTGTGTATGCGCG 360  
Qy  
361 CTCGAGTCTCAGATTGCTACCGAGATGGTGTGTGTACGGCGGAGGTGAAGCTCAT 420  
Db  
361 CTCGAGTCTCAGATTGCTACCGAGATGGTGTGTGTGTACGGCGGAGGTGAAGCTCAT 420  
Qy  
421 AGCGACATCGAAGATCATGATCCCTCTCTGTCCGCTCAGGCGGTAAAGTCATGCTCC 480  
Db  
421 AGCGACATCGAAGATCATGATCCCTCTCTGTCCGCTCAGGCGGTAAAGTCATGCTCC 480  
Qy  
481 CAATTTCTTAGAAAAAGGTATGAACCTTGTGCAAGATGATGGAGATGCGGCTGAGA 540  
Db  
481 CAATTTCTTAGAAAAAGGTATGAACCTTGTGCAAGATGATGGAGATGCGGCTGAGA 540  
Qy  
541 AGGATATGCGCGTGGAGAGTCGCGGTGAAAGAGGCAACAGACTCGAGACCGAAG 600  
Db  
541 AGGATATGCGCGTGGAGAGTCGCGGTGAAAGAGGCAACAGACTCGAGACCGAAG 600  
Qy  
601 GAGTCGATGTAAGGATGGGTGCGTCAAGTACTCTGCAAGTCAAGTCTTTCAGGAT 660  
Db  
601 GAGTCGATGTAAGGATGGGTGCGTCAAGTACTCTGCAAGTCAAGTCTTTCAGGAT 660  
Qy  
661 TTGACTATAGAGGAGCTCGCTCAGAACAGCAATGAGCTATGTCGCTTTTGTGCG 720  
Db  
661 TTGACTATAGAGGAGCTCGCTCAGAACAGCAATGAGCTATGTCGCTTTTGTGCG 720  
Qy  
721 GACTTACCGATGGGTTGCTCCTACCTTGGACTCGTTCAAGGCTATCATGTGGATTTG 780  
Db  
721 GACTTACCGATGGGTTGCTCCTACCTTGGACTCGTTCAAGGCTATCATGTGGATTTG 780  
Qy  
781 TACCTTACTTCGAACTATGAACGGAGATGATGATACCTTTGACTCAGGATTAGCAG 840  
Db  
781 TACCTTACTTCGAACTATGAACGGAGATGATGATACCTTTGACTCAGGATTAGCAG 840  
Qy  
841 TTTCGACGAGTTGGGATCAGCTTATGAGCAAAAGACGCGGCTGCTGGCTCAG 900  
Db  
841 TTTCGACGAGTTGGGATCAGCTTATGAGCAAAAGACGCGGCTGCTGGCTCAG 900  
Qy  
901 CTTCCGATCAGGCTGTTGATAAAGATGTTCAAGTTCGGATATCCTTAAGTCTCCTAG 960  
Db  
901 CTTCCGATCAGGCTGTTGATAAAGATGTTCAAGTTCGGATATCCTTAAGTCTCCTAG 960  
Qy  
961 TGAGAGCAAAACATCGCCCAACCTGCTGAATCTCAAAAGCTGTCGATGAGGAGTAC 1020  
Db  
961 TGAGAGCAAAACATCGCCCAACCTGCTGAATCTCAAAAGCTGTCGATGAGGAGTAC 1020  
Qy  
1021 TCGCTCAGATCAGTAACCTGTTATGCTGGATATGAACCTTTCGACAGCTTGCAT 1080  
Db  
1021 TCGCTCAGATCAGTAACCTGTTATGCTGGATATGAACCTTTCGACAGCTTGCAT 1080  
Qy  
1081 GGATGTTTCAACGACTCTCAGAGCAAAAGCGCTTCAGGATAAACTTCGAGAGAAATTT 1140  
Db  
1081 GGATGTTTCAACGACTCTCAGAGCAAAAGCGCTTCAGGATAAACTTCGAGAGAAATTT 1140  
Qy  
1141 GTCAGATCGACACGATATGCTACGCTAGACGAACCTTAATGCTGCTTATCTCGAAG 1200  
Db  
1141 GTCAGATCGACACGATATGCTACGCTAGACGAACCTTAATGCTGCTTATCTCGAAG 1200  
Qy  
1201 CGTTTGTAAAGAGTCTCTCGTCTAGACCTCTAGTCCGTATGCTAAACCGTGAATGCT 1260  
Db  
1201 CGTTTGTAAAGAGTCTCTCTCGTCTAGACCTCTAGTCCGTATGCTAAACCGTGAATGCT 1260  
Qy  
1261 TAAAGGATGAAGACTTCAATCCACTTCCGAGCTGTCATTTGTCGAGATGGTCCGCTCA 1320  
Db  
1261 TAAAGGATGAAGACTTCAATCCACTTCCGAGCTGTCATTTGTCGAGATGGTCCGCTCA 1320

1321 TCAACGAGGTCCGGATCAGAAAGGAACGATGGTCAATGCTTCGTTGTTCAACATCAATC 1380  
Db  
1321 TCAACGAGGTCCGGATCAGAAAGGAACGATGGTCAATGCTTCGTTGTTCAACATCAATC 1380  
Qy  
1381 GTTCAAAGTTCATTTATGGAAGATGCAAGAAATTCAGACCGAGAGGTGGCTTGAGG 1440  
Db  
1381 GTTCAAAGTTCATTTATGGAAGATGCAAGAAATTCAGACCGAGAGGTGGCTTGAGG 1440  
Qy  
1441 ACCTAACAGACTCGCTCAACAGTATTGAAGCAACCTATGGACACCAAGGCGCTTATCT 1500  
Db  
1441 ACCTAACAGACTCGCTCAACAGTATTGAAGCAACCTATGGACACCAAGGCGCTTATCT 1500  
Qy  
1501 CTGACACCAAGCTTCTGCTTGGTGGCGATTTGCTGTGCGCCAGATGAAGGCTTCTTGT 1560  
Db  
1501 CTGACACCAAGCTTCTGCTTGGTGGCGATTTGCTGTGCGCCAGATGAAGGCTTCTTGT 1560  
Qy  
1561 TTGTCACTCTCGCTCGGTCAGTTCGAGCCCATCATCTCTCATCCAGATACGAGCACA 1620  
Db  
1561 TTGTCACTCTCGCTCGGTCAGTTCGAGCCCATCATCTCTCATCCAGATACGAGCACA 1620  
Qy  
1621 TCACCTTGATCATTTCCCGTCTCGAATCGTTGGTAGAGAAAGGAGGTACCAAGATGC 1680  
Db  
1621 TCACCTTGATCATTTCCCGTCTCGAATCGTTGGTAGAGAAAGGAGGTACCAAGATGC 1680  
Qy  
1681 GTTTCAGGTCAGACCGGTCGATGAGTTGATTTCTTATATGTTAAGAGAGTTCTATAT 1740  
Db  
1681 GTTTCAGGTCAGACCGGTCGATGAGTTGATTTCTTATATGTTAAGAGAGTTCTATAT 1740  
Qy  
1741 CTGAGATGTTGACTAGGACAAATGCTTCTTTGATCGAATTTGTTCTCATACCCGGC 1800  
Db  
1741 CTGAGATGTTGACTAGGACAAATGCTTCTTTGATCGAATTTGTTCTCATACCCGGC 1800  
Qy  
1801 AGGCGCTATGACTCTTACGTCGCTATGCTCGCTCTGACTCTCTCTTACCCCTATAT 1860  
Db  
1801 AGGCGCTATGACTCTTACGTCGCTATGCTCGCTCTGACTCTCTCTTACCCCTATAT 1860  
Qy  
1861 TATTCATCCGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGG 1920  
Db  
1861 TATTCATCCGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGG 1920  
Qy  
1921 CGTCCGAAATTC 1932  
Db  
1921 CGTCCGAAATTC 1932

## RESULT 2

US-09-518-386B-4  
; Sequence 4, Application US/09518386B  
; Patent No. 6365386  
; GENERAL INFORMATION:  
; APPLICANT: HOSHINO, Tatsuo  
; APPLICANT: OJIMA, Kazuyuki  
; APPLICANT: SETOGUCHI, Yutaka  
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
; CURRENT APPLICATION NUMBER: US/09/518,386B  
; CURRENT FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: EP 99104668.1  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EP 00101666.5  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 3969  
; TYPE: DNA  
; ORGANISM: Phaffia rhodozyma  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (517)..(518)  
; NAME/KEY: intron  
; LOCATION: (784)..(898)  
; NAME/KEY: intron

LOCATION: (1016)..(1087)  
NAME/KEY: intron  
LOCATION: (1180)..(1302)  
NAME/KEY: intron  
LOCATION: (1518)..(1600)  
NAME/KEY: intron  
LOCATION: (1635)..(1723)  
NAME/KEY: intron  
LOCATION: (1867)..(1939)  
NAME/KEY: intron  
LOCATION: (2000)..(2081)  
NAME/KEY: intron  
LOCATION: (2182)..(2257)  
NAME/KEY: intron  
LOCATION: (2355)..(2431)  
NAME/KEY: intron  
LOCATION: (2543)..(2618)  
NAME/KEY: intron  
LOCATION: (2653)..(2742)  
NAME/KEY: intron  
LOCATION: (2815)..(2962)  
NAME/KEY: intron  
LOCATION: (3051)..(3113)  
NAME/KEY: intron  
LOCATION: (3172)..(3247)  
NAME/KEY: intron  
LOCATION: (3322)..(3398)  
NAME/KEY: intron  
LOCATION: (3424)..(3513)  
NAME/KEY: polyA site  
LOCATION: (3865)..(3866)  
NAME/KEY: intron  
LOCATION: (653)..(734)  
US-09-518-386B-4

Query Match 18.3%; Score 353.2; DB 4; Length 3969;  
Best Local Similarity 99.2%; Pred. No. 2.1e-97;  
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1514 TTGCTTTGGTGGCGATTGCTGTCGCGAGATGAAGCCCTTCTTTGTGCACTCTCCG 1573  
DB 3508 TCGCAGTGTGGCGATTGCTGTCGCGAGATGAAGCCCTTCTTTGTGCACTCTCCG 3567  
QY 1574 TCGGTCGAGTTCGAGCCCATCATCTCTCATCCAGATGACGACATCACCCTGATCAT 1633  
DB 3568 TCGGTCGAGTTCGAGCCCATCATCTCTCATCCAGATGACGACATCACCCTGATCAT 3627  
QY 1634 TTCCCGTCTCGAATCGTTGGTAGAGAGAGAGGGGTACAGATGCGTTTGCAGGTCAA 1693  
DB 3628 TTCCCGTCTCGAATCGTTGGTAGAGAGAGAGGGGTACAGATGCGTTTGCAGGTCAA 3687  
QY 1694 GCGGTGCAATGAGTTGATTCTTCATATGTTTAAGAGAGTTCTATATCTGAGAAATGTGTG 1753  
DB 3688 GCGGTGCAATGAGTTGATTCTTCATATGTTTAAGAGAGTTCTATATCTGAGAAATGTGTG 3747  
QY 1754 ACTAGACAAATGCTTCTTTGTATGATTTGTTCTCATACCGGGGAGGCGCTATGACT 1813  
DB 3748 ACTAGACAAATGCTTCTTTGTATGATTTGTTCTCATACCGGGGAGGCGCTATGACT 3807  
QY 1814 TCTACGTCGCTATCGTCTGCTGACTCTCTTCTTACCTATATATTATTCATCCG 1871  
DB 3808 TCTACGTCGCTATCGTCTGCTGACTCTCTTCTTACCTATATATTATTCATCCG 3865

RESULT 3  
US-08-277-031B-4  
; Sequence 4, Application US/08277031B  
; Patent No. 6620593  
; GENERAL INFORMATION:  
; APPLICANT: Hayashi, Koji  
; APPLICANT: Sakaki, Toshiyuki  
; APPLICANT: Yabusaki, Yoshiyasu  
; APPLICANT: Komai, Koichiro

APPLICANT: Kaneko, Hideo  
APPLICANT: Nakatsuka, Iwao  
TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF  
TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING  
TITLE OF INVENTION: HUMAN CYTOCHROME P450  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5inch, 1.44MB  
COMPUTER: IBM PC  
OPERATING SYSTEM: Dos 5.0  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/277,031B  
FILING DATE: 19-JULY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP-201120/1993  
APPLICATION NUMBER: JP-180246/1993  
APPLICATION NUMBER: JP-208279/1993  
FILING DATE: 20-07-1993  
FILING DATE: 21-07-1993  
FILING DATE: 30-07-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Raymond C. Stewart  
REGISTRATION NUMBER: 21,066  
REFERENCE/DOCKET NUMBER: 20-3530P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1512  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-277-031B-4

Query Match 2.5%; Score 48.4; DB 4; Length 1512;  
Best Local Similarity 50.7%; Pred. No. 0.00023;  
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
QY 997 AAAAGCTGTCCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056  
DB 863 AAGCTCTGTCCATCTGGAGCTCGTGGCCCAATCAATTAATTTATTTTGTGGCTATG 922  
QY 1057 AAACTTCTTCGACAGTCTTTGACATGATGTTTCCGAGTCTCAGAAAGACAAGCCGTTTC 1116  
DB 923 AAACACGAGGAGTGTCTCTCTCTTATGATGATGAACTGGCCCACTCACCCCTGATGCC 982  
QY 1117 AGGATAACTTCGAGAAAGAAATTTGTGATGATGCGACAGGATATGCTTACGCTAGACGAAC 1176  
DB 983 AGCAGAACTTCGAGGAGAAATTTGATGAGTTTACCCATTAAGGACCAACCCACTATG 1042  
QY 1177 TTAATGCGGTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGTTGAGCC 1230  
DB 1043 ATACTGTGTACAGATGGAGTATCTTGACATGTTGGTGAATGAAACGCTCAGATTATTC 1102  
QY 1231 CTCCTAGTCCGATCTAACCCTGATGCTTAAAGGATGAGACTTCA 1278  
DB 1103 CAATTGCTATGAGACTTGAGAGGGTCTCGAAAAAAGATGTTGAGATCA 1150

RESULT 4  
US-09-023-655-1062  
; Sequence 1062, Application US/09023655

```
/ Patent No. 6607879
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Seilhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
/ TITLE OF INVENTION: EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HERewith
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1062:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2059 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g181373
/ US-09-023-655-1062

Query Match 2.5%; Score 48.4; DB 4; Length 2059;
Best Local Similarity 50.7%; Pred. No. 0.00028;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGCGATGAGGAGGACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATG 1056
Db 953 AAGCTCTGCGATGAGGAGGACTCGCTGAGCAATCAATATCTTTATTTGCTGGCTATG 1012
QY 1057 AAATCTTTCGACGCTTTGACATGATGTTTCCAGGACTCTCAGAGACAAAGCCGTT 1116
Db 1013 AAACACGAGCAGTGTTCTCTCTTCATATGATGAACTGGCCACTCACCTCTGATGCC 1072
QY 1117 AGGATAAATTCGAGAAGAAATTTGTCAGATCGACACGAGATGCTTACGCTAGACGAAC 1176
Db 1073 AGCAGAAATCTGAGGAGGAAATGATGCACTTTTACCAATAAGGCACCAACCCACCTATG 1132
QY 1177 TTAATCGGTTGC-----CTTATCTCGAGCGTTTGAAGGAGTCTCTTCTGCTAGACC 1230
Db 1133 ATACTGTGCTACAGATGAGGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTTCC 1192
QY 1231 CTCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1193 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGAGATGTTGAGATCA 1240

RESULT 5
US-09-144-367-1
```

```
/ Sequence 1, Application US/09144367
/ Patent No. 6432639
/ GENERAL INFORMATION:
/ APPLICANT: Lichter, Jay
/ APPLICANT: Guido, Marco
/ TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
/ FILE REFERENCE: SEQ-12P
/ CURRENT APPLICATION NUMBER: US/09/144,367
/ CURRENT FILING DATE: 1998-08-31
/ PRIOR FILING DATE: 1998-08-31
/ PRIOR FILING DATE: 1997-09-10
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 2759
/ TYPE: DNA
/ ORGANISM: H. sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (70)...(1581)
/ OTHER INFORMATION: Human CYP3A4 cDNA reference sequence
/ US-09-144-367-1

Query Match 2.5%; Score 48.4; DB 4; Length 2759;
Best Local Similarity 50.7%; Pred. No. 0.00034;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGGAGGACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATG 1056
Db 932 AAGCTCTGCGATGAGGAGGACTCGTGCCCAATCAATATCTTTATTTGCTGGCTATG 991
QY 1057 AAATCTTTCGACAGTCTTGACATGGATGTTTCCGAGCTCTCAGAGACAAAGCCGTT 1116
Db 992 AAACACGAGCAGTGTTCTCTCTCTTATGATGATGAACTGGCCACTCACCTCTGATGCC 1051
QY 1117 AGGATAAATTCGAGAAGAAATTTGTCAGATCGACACGAGATGCTTACGCTAGACGAAC 1176
Db 1052 AGCAGAAATCTGAGGAGGAAATGATGCACTTTTACCAATAAGGCACCAACCCACCTATG 1111
QY 1177 TTAATCGGTTGC-----CTTATCTCGAAGCGTTTGAAGGAGTCTCTTCTGCTAGACC 1230
Db 1112 ATACTGTGCTACAGATGAGGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTTCC 1171
QY 1231 CTCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1172 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGAGATGTTGAGATCA 1219

RESULT 6
US-09-716-129-12
/ Sequence 12, Application US/09716129
/ Patent No. 6632920
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 36 Human Secreted Proteins
/ FILE REFERENCE: P2025P1
/ CURRENT APPLICATION NUMBER: US/09/716,129
/ CURRENT FILING DATE: 2000-11-17
/ PRIOR APPLICATION NUMBER: 60/076,053
/ PRIOR FILING DATE: 1998-02-26
/ PRIOR APPLICATION NUMBER: 60/076,057
/ PRIOR FILING DATE: 1998-02-26
/ PRIOR APPLICATION NUMBER: 60/076,052
/ PRIOR FILING DATE: 1998-02-26
/ PRIOR APPLICATION NUMBER: 60/076,054
/ PRIOR FILING DATE: 1998-02-26
/ PRIOR APPLICATION NUMBER: 60/076,051
/ PRIOR FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 186
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 2160
/ TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-716-129-12

Query Match          2.5%   Score 47.4; DB 4; Length 2160;
Pred Local Similarity 70.8%; Pred. No. 0.00059;
Matches 63; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY      1826 ATCTGCGTCTGGACTCTCTTCATCCCTATATATTCCATCCGAATAAAAAAAAAA 1885
DB      2069 ATTGTGCGGTGTACTAGAACTTGACCTTAATAAAAATCCCAAAAGTTAAAAAAA 2128
QY      1886 AAAAAAAAAAAAAAAAAAGCGCGCTCGAG 1914
DB      2129 AAAAAAAAAAAAAAAAAAGCGCGCTCGG 2157

RESULT 7
US-09-148-545-106
; Sequence 106, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,580
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
```

```

; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,514
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 1712

```

```

Query Match      2.4%; Score 46.8; DB 4; Length 1712;
Best Local Similarity 85.0%; Pred. NO. 0.00077;
Matches 51; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1854 TATATATTATTCATCCGAAAAAAGGCGCGCTCGA 1913
    |||||
Db 1645 TAAATATATTTTCTACAAAAAAGGCGCGCTCGA 1704

RESULT 8
US-09-148-545-105
; Sequence 105, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; EARLIER FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23

```

```

; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864

```

```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 1822

```

Query Match 2.4%; Score 46.8; DB 4; Length 1822;

Best Local Similarity 85.0%; Pred. No. 0.0008;

Matches 51; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1854 TATATATATTCATCCGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGA 1913

Db 1755 TAAATATATTTTGTACAAAAAAAAAAAAAAAAAAAAAAAAAGSGCGCTCGA 1814

RESULT 9  
US-08-948-564-15  
; Sequence 15, Application US/08948564  
; Patent No. 6121512  
; GENERAL INFORMATION:  
; APPLICANT: Siminazky, Balazs  
; APPLICANT: Dewey, Ralph E.  
; APPLICANT: Corbin, Frederick T.  
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and  
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Virginia C. Bennett  
; STREET: PO Box 37428  
; CITY: Raleigh  
; STATE: No. 6121512th Carolina  
; COUNTRY: USA  
; ZIP: 27627  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/948,564  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Virginia C.  
; REGISTRATION NUMBER: 37,092  
; REFERENCE/DOCKET NUMBER: 5051-409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-854-1400  
; TELEFAX: 919-854-1401  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1831 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 20..1747  
; US-08-948-564-15  
Query Match 2.3%; Score 45; DB 3; Length 1831;  
Best Local Similarity 52.8%; Pred. No. 0.0029;  
Matches 121; Conservative 0; Mismatches 105; Indels 3; Gaps 1;  
QY 1008 GATGAGGAGTACTCGCTCAGATCAGTACCTGTTATTGTTGATGATGAACTTCTTCG 1067  
Db |||||  
QY 1037 GATCGTCAGTTGAGGATGATTAATGACAATGCTTATTGCGGTCATGAAACACGCT 1096  
Db |||||  
QY 1068 ACAGTCTTGACATGATGTTTACCGACTCTCAGAGACAAAGCGTTTCAGGATAAATT 1127  
Db |||||  
QY 1097 GCAGTTCTTACTTGGCAGCTTTTCTCTAGTCAAATCTAGCAAAATGAAGAGCT 1156  
Db |||||  
QY 1128 CGAGAGA---AATTGTCAGATCGACACGATATGCTTACCTAGACGAACTTAATGCG 1184  
Db |||||  
QY 1157 CAAGCAGAGTAGATTGTTGTTGGGTACGGGAGGCCAACTTTTGAATCACTTAAGGAA 1216  
Db |||||  
QY 1185 TTGCTTATCTCGAACGTTTGTAAAGGAGTCTTCTGCTAGACCCCTC 1233  
Db |||||  
QY 1217 TTGAGTACATTAAGATTGTTGAGGAGCTCTTCGTTATACCCCTC 1265  
Db |||||  
RESULT 10  
US-09-221-013A-7  
; Sequence 7, Application US/09221013A  
; Patent No. 6495740

; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: PCT/AU97/00402  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 3828  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (239)..(3490)  
; US-09-221-013A-7  
Query Match 2.3%; Score 44; DB 4; Length 3828;  
Best Local Similarity 83.3%; Pred. No. 0.0092;  
Matches 50; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1854 TATATATTATTCATCCGAAAAAAGGCGCGCTCGA 1913  
Db |||||  
QY 3746 TACACACTTGTAAACAAAAAAGGCGCGCTCGA 3805  
Db |||||  
RESULT 11  
US-09-144-367-8  
; Sequence 8, Application US/09144367  
; Patent No. 6432639  
; GENERAL INFORMATION:  
; APPLICANT: Lichter, Jay  
; APPLICANT: Guido, Marco  
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4  
; FILE REFERENCE: SEQ-12P  
; CURRENT APPLICATION NUMBER: US/09/144,367  
; CURRENT FILING DATE: 1998-08-31  
; PRIOR APPLICATION NUMBER: 60/059,612  
; PRIOR FILING DATE: 1997-09-10  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 352  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: Other  
; LOCATION: (0)...(0)  
; US-09-144-367-8  
Query Match 2.3%; Score 43.6; DB 4; Length 352;  
Best Local Similarity 57.2%; Pred. No. 0.0027;  
Matches 79; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
QY 1002 CTGTCCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGATGAACT 1061  
Db |||||  
QY 123 CTGTCCGATCTGGAGCTCGTCCCAATCAATTATCTTTTGTGCTATGAAAC 182  
Db |||||  
QY 1062 TCTTCGACAGTCTTGACATGATGTTTACCGACTCTCAGAGACAAAGCGTTTCAGAT 1121  
Db |||||  
QY 183 ACAGCAGTGTCTCTCTCTTATGTAATGCACTGCCACTCACCTGATGTCAGCAG 242  
Db |||||  
QY 1122 AAACCTTCGAGAAATT 1139  
Db |||||  
QY 243 AAACCTGCAGGAAATT 260  
Db |||||

```

RESULT 12
US-09-372-422A-9
; Sequence 9, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372.422A
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (142)...(1018)
US-09-372-422A-9
Query Match      2.2%; Score 42.8; DB 4; Length 1333;
Best Local Similarity 87.0%; Pred. No. 0.011; 7; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1855 ATATATTATTCATCCGAGAAAAAAGAAAAAAGAAAAAAGCGCGC 1908
DB 1280 ATCGATTATTCCTCCCAAAAAAAGAAAAAAGAAAAAAGCGCGC 1333

RESULT 13
US-09-370-807-7
; Sequence 7, Application US/09370807
; Patent No. 6297034
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/370,807
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/096,225
; EARLIER FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-370-807-7
Query Match      2.2%; Score 42.8; DB 3; Length 2407;
Best Local Similarity 80.6%; Pred. No. 0.016; 12; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1840 CTCCTCTTACCTATATTTCCATCCGAAAAAAGAAAAAAGAAAAA 1899
DB 1939 CTATCATCTTAGCGTTAATCTCTCCATGAAAAAAGAAAAAAG 1998

QY 1900 AA 1901
DB 1999 AA 2000

RESULT 14
US-09-921-259-7
; Sequence 7, Application US/09921259

```

```

; Patent No. 6465234
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/921,259
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/096,225
; PRIOR FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-921-259-7
Query Match      2.2%; Score 42.8; DB 4; Length 2407;
Best Local Similarity 80.6%; Pred. No. 0.016; 12; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1840 CTCCTCTTACCTATATTTCCATCCGAAAAAAGAAAAAAGAAAAA 1899
DB 1939 CTATCATCTTAGCGTTAATCTCTCCATGAAAAAAGAAAAAAG 1998

QY 1900 AA 1901
DB 1999 AA 2000

RESULT 15
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872337
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolffe, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872337el
; TITLE OF INVENTION: Sequences and Antibodies Thereof
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:

```



```

; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match      2.2%; Score 42.6; DB 2; Length 246240;
Best Local Similarity 67.4%; Pred. No. 0.34;
Matches 50; Conservative 29; Indels 0; Gaps 0;

QY 1813 TTCTACGTCGTCTATCGCTCGCTCTGACTCTCTTACCCCTATATATTCCATCCGA 1872
DB 185758 TCCTACTGCCCATAAAGACAGGCAAACTCCTGGGCTTCCCTGTATATGCTTCCCGCAA 185817

QY 1873 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1901
DB 185818 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 185846

```

Search completed: March 31, 2004, 10:41:39  
Job time : 177 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 13:24:50 ; Search time 742 Seconds

(without alignments)

9758.296 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932

Sequence: 1 gaattcgccagggccacc.....agccggctctgcgaattc 1932

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2470632 seqs, 187387510 residues

Total number of hits satisfying chosen parameters: 4941264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubnpa/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubnpa/US09A\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubnpa/US09B\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubnpa/US09C\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubnpa/US10A\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubnpa/US10B\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubnpa/US10C\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1932	100.0	1932	14	US-10-066-007-2
2	353.2	18.3	3969	14	Sequence 2, Appli
3	49.8	2.6	467	12	Sequence 4, Appli
4	49.4	2.6	1501	15	Sequence 20949, A
5	49.2	2.5	368	12	Sequence 392, App
6	48.4	2.5	657	12	Sequence 1793, Ap
7	48.4	2.5	1512	15	Sequence 25496, A
8	48.4	2.5	2011	9	Sequence 55, Appl
9	48.4	2.5	2759	14	Sequence 1586, Ap
10	48.4	2.5	2768	14	Sequence 1, Appli
11	48.2	2.5	2849	9	Sequence 14, Appl
12	48.2	2.5	312	9	Sequence 2110, Ap
13	47.6	2.5	3059	9	Sequence 4611, Ap
14	47.4	2.5	628	12	Sequence 268, App
15	47.4	2.5	1410	12	Sequence 71313, A
					Sequence 29, Appl

C 16	47.2	2.4	1228	12	US-10-424-599-32438	Sequence 32438, A
C 17	47	2.4	218	9	US-09-933-797-162	Sequence 162, App
C 18	46.8	2.4	685	12	US-10-424-599-8656	Sequence 8556, Ap
C 19	46.8	2.4	709	12	US-10-424-599-75080	Sequence 75080, A
20	46.8	2.4	1151	12	US-10-424-599-23175	Sequence 23175, A
21	46.8	2.4	1712	9	US-09-981-876-106	Sequence 106, App
22	46.8	2.4	1712	10	US-09-148-545-106	Sequence 106, App
23	46.8	2.4	1822	9	US-09-981-876-105	Sequence 105, App
24	46.8	2.4	1822	10	US-09-148-545-105	Sequence 105, App
C 25	46.6	2.4	1091	12	US-10-424-599-72966	Sequence 72966, A
C 26	46.4	2.4	1419	12	US-10-424-599-106636	Sequence 9744, A
C 27	46.2	2.4	672	12	US-10-424-599-99744	Sequence 99744, A
C 28	46.2	2.4	710	12	US-10-424-599-15121	Sequence 15121, A
C 29	46.2	2.4	1194	12	US-10-424-599-29930	Sequence 29930, A
C 30	46	2.4	266	12	US-10-424-599-64827	Sequence 64827, A
C 31	46	2.4	414	12	US-10-424-599-94220	Sequence 94220, A
C 32	46	2.4	940	12	US-10-424-599-11262	Sequence 11262, A
C 33	45.8	2.4	580	12	US-10-424-599-54757	Sequence 54757, App
C 34	45.8	2.4	592	14	US-10-066-543-950	Sequence 54758, A
C 35	45.8	2.4	991	12	US-10-424-599-54758	Sequence 54758, A
C 36	45.8	2.4	1498	12	US-10-424-599-138418	Sequence 138418, A
C 37	45.8	2.4	2301	12	US-10-424-599-142657	Sequence 142657, A
C 38	45.6	2.4	447	12	US-10-424-599-52047	Sequence 52047, A
C 39	45.6	2.4	583	12	US-10-424-599-137273	Sequence 137273, A
C 40	45.6	2.4	705	12	US-10-424-599-118266	Sequence 118266, A
C 41	45.6	2.4	733	12	US-10-424-599-97712	Sequence 97712, A
C 42	45.4	2.3	558	12	US-10-424-599-67477	Sequence 67477, A
C 43	45.4	2.3	662	12	US-10-424-599-80087	Sequence 80087, A
C 44	45.4	2.3	767	12	US-10-424-599-43674	Sequence 43674, A
C 45	45.4	2.3	1353	15	US-10-369-493-43674	Sequence 43674, A

## ALIGNMENTS

### RESULT 1

US-10-066-007-2  
Sequence 2, Application US/10066007  
Publication No. US20030077691A1  
GENERAL INFORMATION:

APPLICANT: HOSHINO, Tatsuo  
APPLICANT: OJIMA, Kazuyuki  
APPLICANT: SETOGUCHI, Yutaka  
TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
CURRENT APPLICATION NUMBER: US/10/066,007  
CURRENT FILING DATE: 2001-02-01  
PRIOR APPLICATION NUMBER: US/09/518,386  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: EP 99104668.1  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: EP 00101666.6  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 1932  
TYPE: DNA  
ORGANISM: Phaffia rhodozyma  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (33)..(1706)  
NAME/KEY: polyA site  
LOCATION: (1871)  
NAME/KEY: mRNA  
LOCATION: (14)..(1891)  
US-10-066-007-2

Query Match 100.0%; Score 1932; DB 14; Length 1932;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1932; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GAATTCGGCAGGAGCCACTTCTCCATATGTTTCATCTTGTTGTCACAGTG 60

Db	1	GAATTCGGCAGGAGCCACTCTTTCTCCATATGTTCTCATCTTGTTGCTTGTCTCACAGGTG	60
Qy	61	CTTTAGGCTGGCTGCTTCTCATGGGACCATAGCGTTCCTTACGTCTTTACCTCGCTC	120
Db	61	CTTTAGGCTGGCTGCTTCTCATGGGACCATAGCGTTCCTTACGTCTTTACCTCGCTC	120
Qy	121	CGAGGGGATCTTCACTGTATAACTTTCAGGGCCGGAATCATACCAACTACTTTACAGCA	180
Db	121	CGAGGGGATCTTCACTGTATAACTTTCAGGGCCGGAATCATACCAACTACTTTACAGCA	180
Qy	181	ATTTTTTAGACATCTCTCAGCTCGTACAGGTGAAGAGCATCGAAGTACAGAGAAAAT	240
Db	181	ATTTTTTAGACATCTCTCAGCTCGTACAGGTGAAGAGCATCGAAGTACAGAGAAAAT	240
Qy	241	ACGGAAGCACCCCTCCGGTTTGTCTGGGATCGCTGGAGCACCCGCTTTGAACTCGACGATC	300
Db	241	ACGGAAGCACCCCTCCGGTTTGTCTGGGATCGCTGGAGCACCCGCTTTGAACTCGACGATC	300
Qy	301	CGAAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCGAAACCTGGTATGGCCG	360
Db	301	CGAAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCGAAACCTGGTATGGCCG	360
Qy	361	CTCGAGTCTCAGAAATGTCTACCGGAGATGGTGTGTACGGCGGAAGTGAAGTCTATA	420
Db	361	CTCGAGTCTCAGAAATGTCTACCGGAGATGGTGTGTACGGCGGAAGTGAAGTCTATA	420
Qy	421	AGGCATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCCGTTAAGTCGATGGTCC	480
Db	421	AGGCATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCCGTTAAGTCGATGGTCC	480
Qy	481	CAATTTCTTAGAAAAAGGTATGGAATCTCTGCACAAGATGATGGAGGATGCGCTGAGA	540
Db	481	CAATTTCTTAGAAAAAGGTATGGAATCTCTGCACAAGATGATGGAGGATGCGCTGAGA	540
Qy	541	AGGATATGGCCGTGGGAGGTCGGCCCGGTGAAAAGAGGCAACAGACTCGAGCCGAAG	600
Db	541	AGGATATGGCCGTGGGAGGTCGGCCCGGTGAAAAGAGGCAACAGACTCGAGCCGAAG	600
Qy	601	GAGTCGATGTAAAGGATTTGGTGGTCGAGCTACTCTGGACGTCATGGCTCTTGCAAGAT	660
Db	601	GAGTCGATGTAAAGGATTTGGTGGTCGAGCTACTCTGGACGTCATGGCTCTTGCAAGAT	660
Qy	661	TTGACTATAAGAGCGACTCGCTCCAGAACCAAGACCAATGAGCTCTATGTCGCTTTTGTG	720
Db	661	TTGACTATAAGAGCGACTCGCTCCAGAACCAAGACCAATGAGCTCTATGTCGCTTTTGTG	720
Qy	721	GACTTACCGATGGTTGGTCTCCTACCTTGACTGCTTTCAAGGCTACATGCGGATTTG	780
Db	721	GACTTACCGATGGTTGGTCTCCTACCTTGGACTCGTTCAAGGCTACATGCGGATTTG	780
Qy	781	TACCTTACTCCGAACTATGAAACGGAGACATGAGTACCTTTGACTCAAGGATTAGCAG	840
Db	781	TACCTTACTCCGAACTATGAAACGGAGACATGAGTACCTTTGACTCAAGGATTAGCAG	840
Qy	841	TTTCCCGACGATTTGGGATCGAGCTTATGGAGCAAAAAGACGAGCCGCTGCTGGCTCAG	900
Db	841	TTTCCCGACGATTTGGGATCGAGCTTATGGAGCAAAAAGACGAGCCGCTGCTGGCTCAG	900
Qy	901	CTTCCGATCAGGCTTGTGATAAAAAGGATGTTCAAGGTCGGGATATCCTAAGTCTCCTAG	960
Db	901	CTTCCGATCAGGCTTGTGATAAAAAGGATGTTCAAGGTCGGGATATCCTAAGTCTCCTAG	960
Qy	961	TCAGAGCAAAACATCGCCGCCAACCTGCTGAATCTCAAAAAGCTCTCGATGAGAGGTAC	1020
Db	961	TCAGAGCAAAACATCGCCGCCAACCTGCTGAATCTCAAAAAGCTCTCGATGAGAGGTAC	1020
Qy	1021	TCGCTCAGATCAGTAAACCTGTTATTTGCTGGATATGAACTTCTTCGACAGTCTTGACAT	1080
Db	1021	TCGCTCAGATCAGTAAACCTGTTATTTGCTGGATATGAACTTCTTCGACAGTCTTGACAT	1080
Qy	1081	GGATGTTTTCACCGACTCTCAGAAGCAAAAGCCGTTTCAGGATTAACCTTCGAGAGAAATTT	1140

Db	1081	GGATGTTTCCCGGACTCTCAGAGACAAAGCCGTTCCAGATAAACCCTTCGAGAGAAATTT	1144
Qy	1141	GTCCAGATCGACACGGATATGCCCTACGCTAGACGAACCTTAATGCGTTGCCTTATCTCGAAG	1200
Db	1141	GTCCAGATCGACACGGATATGCCCTACGCTAGACGAACCTTAATGCGTTGCCTTATCTCGAAG	1200
Qy	1201	CGTTTGTGTTAAGGAGTCTCTTCGCTTAGAACCCCTCTAGTCCGTTATGCTAAACCGTGAATGCT	1260
Db	1201	CGTTTGTGTTAAGGAGTCTCTTCGCTTAGAACCCCTCTAGTCCGTTATGCTAAACCGTGAATGCT	1260
Qy	1261	TAAAGGATGGAAGACTTCATCCCACTTCCCGAGCCCTGTCAATTGGTTCGAGATGGGTCCGTCA	1320
Db	1261	TAAAGGATGGAAGACTTCATCCCACTTCCCGAGCCCTGTCAATTGGTTCGAGATGGGTCCGTCA	1320
Qy	1321	TCACGAGGTCGGGATCACGAAAGGAACGATGGTTCATGCTTCGCTTGTTTCAACATCAATC	1380
Db	1321	TCACGAGGTCGGGATCACGAAAGGAACGATGGTTCATGCTTCGCTTGTTTCAACATCAATC	1380
Qy	1381	GTTCAAAGTTCATTTATGGAGAGATGTCAGAGAAATTCAGACCGCAGAGGTGGCTTCGAG	1440
Db	1381	GTTCAAAGTTCATTTATGGAGAGATGTCAGAGAAATTCAGACCGCAGAGGTGGCTTCGAG	1440
Qy	1441	ACGTAAACAGACTCGCTCAACAGATTATGAAGCACCCCTATGGACACCGAGCGAGCTTTATCT	1500
Db	1441	ACGTAAACAGACTCGCTCAACAGATTATGAAGCACCCCTATGGACACCGAGCGAGCTTTATCT	1500
Qy	1501	CTGGACCCAGAGCTTGCTTTGGTTGGCGATTGCTGTCGCCGAGATGAAGGCCCTTCTTGT	1560
Db	1501	CTGGACCCAGAGCTTGCTTTGGTTGGCGATTGCTGTCGCCGAGATGAAGGCCCTTCTTGT	1560
Qy	1561	TTGTCACTCTCCGTCGGGTCCAGTTCGAGGCCATCATCTCTCATCCAGAGTACGAGCACAC	1620
Db	1561	TTGTCACTCTCCGTCGGGTCCAGTTCGAGGCCATCATCTCTCATCCAGAGTACGAGCACAC	1620
Qy	1621	TCACTTGTATCATTTCCGTCCTCGAATCGTTCGTAGAGAGAGAGAGGGGTACCAAGATGC	1680
Db	1621	TCACTTGTATCATTTCCGTCCTCGAATCGTTCGTAGAGAGAGAGAGGGGTACCAAGATGC	1680
Qy	1681	GTTTGCAGGTCAAGCCGGTCGAATGAGTTGATTCCTCATATGTTAAGAGAGTTCTATAT	1740
Db	1681	GTTTGCAGGTCAAGCCGGTCGAATGAGTTGATTCCTCATATGTTAAGAGAGTTCTATAT	1740
Qy	1741	CTGAGAAATGTGACTACGACAAATGCCCTCTTTGTATCGAATTTGTTTCFCAACCCGGGC	1800
Db	1741	CTGAGAAATGTGACTACGACAAATGCCCTCTTTGTATCGAATTTGTTTCFCAACCCGGGC	1800
Qy	1801	AGSCGCTATGACTTCTACGTCGTCTATCGTCGCTTGGACTCTCTTCTTACCCCTATATAT	1860
Db	1801	AGSCGCTATGACTTCTACGTCGTCTATCGTCGCTTGGACTCTCTTCTTACCCCTATATAT	1860
Qy	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGCTCGAGCGCGCT	1920
Db	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGCTCGAGCGCGCT	1920
Qy	1921	CGTGCCGAATTC	1932
Db	1921	CGTGCCGAATTC	1932

```

RESULT 2
US/10-066-007-4
Sequence 4, Application US/10066007
Publication No. US20030077691A1
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: QIIMA, Kazuyuki
APPLICANT: SETOGUCHI, Yutaka
TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
FILE REFERENCE: ASTAXANTHIN SYNTHETASE
CURRENT APPLICATION NUMBER: US/10/066,007
CURRENT FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: US/09/518,386
PRIOR FILING DATE: 2000-03-03

```

Q7 1 5/4 AC TAGGACAAAGCCCTTCTTTTGTATCGATTGTTCATATCCCGCGCGAGCGCCCTATGACCT

	Query Match	Best Local Similarity	74.4%;	Pred. No. 0.00061;	Mismatches 61;	Conservative	0;	Indels	0;	Gaps	0;
Qy	1833	CTCTGGACCTCTCTTCTTCCCTATATATATATCTCATCCGAAAAA	AAAAAAAAAAAAAAAAAAAA	1892							
Db	91	CTCTCGAGATCTCTTGGAAACGATTCAGATGAAAGAGCAAAAAA	AAAAAAAAAAAAAAAAAAAA	32							
Qy	1893	AAAAAAAAAAGCGCGCGCTCGAG	1914								

```

RESULT 8
US-09-880-107-1586
Sequence 1586, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherff, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression P
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880
CURRENT FILING DATE: 2001-06-14
APPLICATION NUMBER: US 60/211,3
PRIOR FILING DATE: 2000-06-14
APPLICATION NUMBER: US 60/237,0
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patent In Ver. 2.1

```

SEQ ID NO 1586  
 LENGTH: 2011  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00003  
 US-09-880-107-1586

Query Match 2.5%; Score 48.4; DB 9; Length 2011;  
 Best Local Similarity 50.7%; Pred. No. 0.0012;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056  
 DB 926 AAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 985  
 QY 1057 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1116  
 DB 986 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1045  
 QY 1117 AGATTAACCTGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1176  
 DB 1046 AGATTAACCTGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1105  
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTGTTAAAGGATCTCTTCTAGAC 1230  
 DB 1106 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1165  
 QY 1231 CTCCTAGTCCGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1278  
 DB 1166 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAGATGTTGAGATCA 1213

RESULT 9  
 US-10-146-575-1  
 ; Sequence 1, Application US/10146575  
 ; Publication No. US20030059800A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lichter, Jay  
 ; APPLICANT: Guido, Marco  
 ; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4  
 ; FILE REFERENCE: SEQ-12P  
 ; CURRENT APPLICATION NUMBER: US/10/146,575  
 ; PRIOR FILING DATE: 2002-05-14  
 ; PRIOR APPLICATION NUMBER: US/09/144,367  
 ; PRIOR FILING DATE: 1998-08-31  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2759  
 ; TYPE: DNA  
 ; ORGANISM: H. sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (70)....(1581)  
 ; OTHER INFORMATION: Human CYP3A4 cDNA reference sequence  
 ; US-10-146-575-1

Query Match 2.5%; Score 48.4; DB 14; Length 2759;  
 Best Local Similarity 50.7%; Pred. No. 0.0015;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056  
 DB 932 AAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 991  
 QY 1057 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1116  
 DB 992 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1051  
 QY 1117 AGATTAACCTGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1176  
 DB 1052 AGCAGAACTGCAGGAGGAAATGTCAGTCTTACCCAAATAGGACACCCACCTATG 1111

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTGTTAAAGGATCTCTTCTAGAC 1230  
 DB 1112 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1171  
 QY 1231 CTCCTAGTCCGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1278  
 DB 1172 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAGATGTTGAGATCA 1219

RESULT 10  
 US-10-268-822-14  
 ; Sequence 14, Application US/10268822  
 ; Publication No. US20030150004A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moore, David  
 ; APPLICANT: Wei, Ping  
 ; APPLICANT: Chua, Steven  
 ; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob  
 ; FILE REFERENCE: P02729US2  
 ; CURRENT APPLICATION NUMBER: US/10/268,822  
 ; CURRENT FILING DATE: 2002-10-10  
 ; PRIOR APPLICATION NUMBER: PCT/US 01/29672  
 ; PRIOR FILING DATE: 2001-09-21  
 ; PRIOR APPLICATION NUMBER: US 10/219,590  
 ; PRIOR FILING DATE: 2002-08-15  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 14  
 ; LENGTH: 2768  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-10-268-822-14

Query Match 2.5%; Score 48.4; DB 14; Length 2768;  
 Best Local Similarity 50.7%; Pred. No. 0.0015;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056  
 DB 967 AAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1026  
 QY 1057 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1116  
 DB 1027 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1086  
 QY 1117 AGATTAACCTGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1176  
 DB 1087 AGCAGAACTGCAGGAGGAAATGTCAGTCTTACCCAAATAGGACACCCACCTATG 1146  
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTGTTAAAGGATCTCTTCTAGAC 1230  
 DB 1147 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1206  
 QY 1231 CTCCTAGTCCGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1278  
 DB 1207 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAGATGTTGAGATCA 1254

RESULT 11  
 US-09-880-107-2110  
 ; Sequence 2110, Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darci T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherf, Uwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT APPLICATION NUMBER: US/09/880,107  
 ; CURRENT FILING DATE: 2001-06-14

SECRET

SECRET

```

US-10-395-463-29
; Sequence 29, Application US/10395463
; Publication No. US20040060079A1
; GENERAL INFORMATION:
; APPLICANT: Tanaka, Hiroshi
; APPLICANT: Kayano, Toshiaki
; APPLICANT: Matsuoka, Makoto
; APPLICANT: Sakamoto, Tomoaki
; APPLICANT: Iwahori, Shuichi
; TITLE OF INVENTION: Method of Controlling Character of Monocotyledon by Modification
; TITLE OF INVENTION: Overexpression of Cytochrome P450 Monooxygenase Gene Involved in
; TITLE OF INVENTION: Biosynthesis and Monocotyledon Modified by the Gene
; FILE REFERENCE: 59150-8022
; CURRENT APPLICATION NUMBER: US/10/395,463
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-276398
; PRIOR FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1410)
; OTHER INFORMATION:
US-10-395-463-29

```

```

Query Match      2.5%; Score 47.4; DB 12; Length 1410;
Best Local Similarity 54.2%; Pred.No. 0.002;
Matches 96; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 997 AAAAGCTGTCGAGTACGAGGAGTACTCGCTCAGATCAGTAACTGTATTGCTGGATATG 1056
Db 791 AGAGCTCAGAGTACGAGATCATTTGACCTGATCATCCCTTATATCTCTGGATATG 850

QY 1057 AAACTTCTTCGACAGTCTTGACATGGATGTTTACCGACTCTCAGAGACAAAGCCGTTTC 1116
Db 851 AAACCATGTCGAGGACCTCGATGATGGCTGCAAGTACCTGTCAGACCATCCCAAAGCTC 910

QY 1117 AGGATAAATTCGAGAAAGAAATTTGTCAGATCGACGCGATATGCTACGCTAGACG 1173
Db 911 TTGAGCAACTCAGGAAGAACATTTTGATATCAGGAAGGTAAAGCCGCCGAGATG 967

```

Search completed: April 6, 2004, 17:58:23  
Job time : 748 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 13:16:26 ; Search time 5193 Seconds  
(without alignments)  
11131.340 Million cell updates/sec

Title: US-10-066-007-2  
Perfect score: 1932  
Sequence: 1 gaattcgacagagccacc.....agcggctcgtgcgaattc 1932

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrt:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	4.6	528	14	CD275123 T143B0028
2	79.2	4.1	474	14	CD275883 T143B0523
3	79.2	4.1	475	14	CD275923 T143B0022
4	75.8	3.9	615	9	AT002896 AT002896

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	69.2	3.5	354	14	CD274181
6	68	3.5	403	14	CD273194
7	61	3.2	487	13	BQ102580
8	59.8	3.0	480	14	CD275427
9	58	3.0	317	14	CD275197
10	56.8	2.9	499	12	BQ705377
11	55.4	2.9	971	13	EX779027
12	53.8	2.8	849	14	CF152584
13	51.2	2.7	568	10	AW600900
14	50.6	2.6	1296	10	AW600865
15	50.4	2.6	875	14	CB200527
16	50.2	2.6	407	13	BU398544
17	50.2	2.6	723	13	BU250947
18	50.2	2.6	951	13	BU246013
19	50	2.6	740	12	BQ207705
20	50	2.6	907	29	CNS015AK
21	49.8	2.6	212	9	AL651909
22	49.8	2.6	303	9	AA907131
23	49.6	2.6	544	14	CF366878
24	49.4	2.6	447	9	AV736322
25	49.2	2.5	387	13	BU779630
26	49	2.5	483	14	CB821156
27	48.8	2.5	455	9	AV744471
28	48.6	2.5	547	12	BI378359
29	48.4	2.5	396	12	BM139856
30	48.4	2.5	480	14	CB157627
31	48.4	2.5	580	14	CB114922
32	48.4	2.5	609	14	CB162338
33	48.4	2.5	641	14	CB154142
34	48.4	2.5	697	9	AV648984
35	48.4	2.5	755	12	BQ206682
36	48.4	2.5	757	12	EG211348
37	48.4	2.5	760	12	EG196339
38	48.4	2.5	763	12	EG206681
39	48.4	2.5	765	12	EG182832
40	48.4	2.5	768	12	BG220280
41	48.4	2.5	770	12	EG197817
42	48.4	2.5	770	12	EG214545
43	48.4	2.5	789	12	EG188595
44	48.4	2.5	790	12	BG206683
45	48.4	2.5	791	12	BG208228

## ALIGNMENTS

RESULT 1	CD275123	528 bp	mrna	linear	EST 01-SEP-2003
LOCUS	T143B00283 (FHIG:B)	Axenic plate culture	Paxillus involutus	CDNA	
DEFINITION	5', mRNA sequence.				
ACCESSION	CD275123				
VERSION	CD275123.1	GI:34387169			
KEYWORDS	EST.				
SOURCE	Paxillus involutus				
ORGANISM	Paxillus involutus				
REFERENCE	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Boletales; Paxillineae; Paxillaceae; Paxillus.				
AUTHORS	Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R., Lundeborg, J., Uhlen, M. and Tunlid, A.				
TITLE	Transcriptional responses of Paxillus involutus and Betula pendula during formation of ectomycorrhizal root tissue				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Johansson, T. Fungal-Host Interaction Group (FHIG) Microbial Ecology, Institution of Ecology Ecology Building, Lund University, SE-223 62 Lund, Sweden Tel: +46 46 222 41 58 Fax: +46 46 222 41 58 Email: tomas.johansson@mbioekol.lu.se PCR Primers FORWARD: P104 (5'-GGGAGCGCGCCATGTGTT-3')				

BACKWARD: P105 (5'-AGTGAGCTCGAATTGGGCC-3')

Seq primer: P104

High quality sequence stop: 528.

Location/Qualifiers

# FEATURES

source

1. .528

/organism="Paxillus involutus"

/mol\_type="rRNA"

/strain="ATCC 200175"

/db\_xref="taxon:71150"

/tissue\_type="Mycelium"

/dev\_stage="25 days of growth after transfer"

/lab\_host="Escherichia coli BM25.8"

/note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This

EST clone is originating from one of three cDNA libraries,

constructed for transcript profiling of the mycorrhizal

interaction between the basidiomycete *Paxillus involutus*

and *Betula pendula* (birch). One library represents the

developed and functional mycorrhizal root tissue

(('FHIG:A) Ectomycorrhiza plate culture'), a second

library represents axenically grown fungus ('(FHIG:B)

Axenic plate culture') and a third library represents

axenically grown plants ('(FHIG:C) Axenic plate culture').

Libraries were analyzed in parallel and 3555 (FHIG:A),

3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)

ESTs of >99bp have been deposited. The cDNA libraries were

constructed from total RNA using the SMART cDNA library

construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)

according to the manufacturer's instructions. Full-length

cDNAs were trimmed by SfiI, fractionated and directionally

ligated into (lambda)Triplex2 arms. The lambda library was

converted to a plasmid library via site-specific

recombination at loxP sites in a Cre+ strain (E. coli

BM25.8). Plasmid clones were randomly collected and

analysed by DNA sequencing using a plasmid-specific

forward primer (P104)."

## ORIGIN

Query Match

Best Local Similarity 4.6%; Score 88; DB 14; Length 528;

Mismatches 244; Conservative 0; Mismatches 195; Indels 9; Gaps 3;

```

QY 1069 CAGCTTTGACATGGATGTTTCCACGACTCTCAGAGACAAGCCGTTTCAGGATAAATTC 1128
DB 88 CACGATAACTTGGGCGCTGCATGAGTTTGCTGCTAGCCCTGGGATACAGAGGAAGTCC 147
QY 1129 GAGAAGAAATTTTCAGATCGACACGATATGCTTACGCTAGACGAACTTAATCGTTGC 1188
DB 148 GCGAGGAACTTCTTCTGTTGACACGGAACGCCCTCGATGAGCAACTCTCGGGCTTC 207
QY 1189 CTTATCTCGAAGGTTTGTAAAGGAGTCTCTTCGTCTAGACCTCTAGTCCGTATGCTA 1248
DB 208 CTTACCTTGACACCGGTGTAAGGAGACTTTGCGTGTACACCTCCCTTTGAGAGACCA 267
QY 1249 ACCGTGAATGCTTAAAGGATGAAGCTTCATCCACTTCCGAGCTCTCATTTGGTCGAG 1308
DB 268 TAGCGG---TCGGCATGAAGACGATATTTCTCCGCTCGAGAGCCACTTCAGCAACAC 324
QY 1309 ATGGGTCGTTATCAACGAGGTCGCGATCAGCAAGGAACGATGCTATGTTTCGTTGT 1368
DB 325 ATGGCGTCGTACACGACGGAATAGGATCAGCAAGGAACGCGCATCTCTATCCCGATCT 384
QY 1369 TCAACATCATGTTCCAAAGTTCAITTTATGAGAGATGTCAGAGAATTCAGACCGGAGA 1428
DB 385 T---GATGAACCGGTGAAAGGGGTTGTGGGGCCCTGACGCCACAGGTTCAACCCGGAAC 441
QY 1429 GTGGGTTGAGGACGTAACAGACTCGCTCAACAGATTTGAAGACCCCTATGACACCCAGG 1488
DB 442 GCTGG---GACGACGTACCTAAAGTGTCTTCCAATATCCAGGTGTTTGGGGCCACATGC 498
QY 1489 CGAGCTTTATCTCTGACACCCAGGCTTG 1516
DB 499 TAAAGTTTCTCTCGCGGGCCGAGACATG 526

```

## RESULT 2

CD275983

LOCUS

DEFINITION

CD275983

VERSION

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CD275983 474 bp mRNA linear EST 01-SEP-2003  
T143B05239F (FHIG:B) Axenic plate culture *Paxillus involutus* cDNA  
5', mRNA sequence.

CD275983 GI:34387929

EST.

*Paxillus involutus*

*Paxillus involutus*

Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

Boletales; Paxillaceae; Paxillaceae; *Paxillus*.

1 (bases 1 to 474)

Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,

Lundberg, J., Uhlen, M. and Tunlid, A.

Transcriptional responses of *Paxillus involutus* and *Betula pendula*

during formation of ectomycorrhizal root tissue

Unpublished (2003)

CONTACT: Johansson, T.

Fungal-Hest Interaction Group (FHIG)

Microbial Ecology, Institution of Ecology

Ecology Building, Lund University, SE-223 62 Lund, Sweden

Tel: +46 46 222 45 49

Fax: +46 46 222 41 58

Email: tomas.johansson@mbioekol.lu.se

PCR Primers

FORWARD: P104 (5'-GGGAGCGCGCATTTGTT-3')

BACKWARD: P105 (5'-AGTGAGCTCGAATTGGGCC-3')

Seq primer: P104

High quality sequence stop: 474.

Location/Qualifiers

1. .474

source

/organism="Paxillus involutus"

/mol\_type="mRNA"

/strain="ATCC 200175"

/db\_xref="taxon:71150"

/tissue\_type="Mycelium"

/dev\_stage="25 days of growth after transfer"

/lab\_host="Escherichia coli BM25.8"

/clone\_lib="(FHIG:B) Axenic plate culture"

/note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This

EST clone is originating from one of three cDNA libraries,

constructed for transcript profiling of the mycorrhizal

interaction between the basidiomycete *Paxillus involutus*

and *Betula pendula* (birch). One library represents the

developed and functional mycorrhizal root tissue

('(FHIG:A) Ectomycorrhiza plate culture'), a second

library represents axenically grown fungus ('(FHIG:B)

Axenic plate culture') and a third library represents

axenically grown plants ('(FHIG:C) Axenic plate culture').

Libraries were analyzed in parallel and 3555 (FHIG:A),

3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)

ESTs of >99bp have been deposited. The cDNA libraries were

constructed from total RNA using the SMART cDNA library

construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)

according to the manufacturer's instructions. Full-length

cDNAs were trimmed by SfiI, fractionated and directionally

ligated into (lambda)Triplex2 arms. The lambda library was

converted to a plasmid library via site-specific

recombination at loxP sites in a Cre+ strain (E. coli

BM25.8). Plasmid clones were randomly collected and

analysed by DNA sequencing using a plasmid-specific

forward primer (P104)."

## ORIGIN

Query Match

Best Local Similarity 4.1%; Score 79.2; DB 14; Length 474;

Mismatches 203; Conservative 0; Mismatches 163; Indels 6; Gaps 2;

QY 1069 CAGCTTTGACATGGATGTTTCCACGACTCTCAGAGACAAGCCGTTTCAGGATAAATTC 1128

```

Db      79  CAGCGATTAACCTTGGCGCTGATGATGTTGTCGCTAGCCCTGGGATACAGAGAACTCC 138
QY      1129 GAGAGAAATTTTCAGATCGACACGGATATGCTTACGCTAGACGAATTAATGCGTTGC 1188
Db      139  GGGAGAACTCTTCTCTGTGACACGGAACGCCCTGATGGAGAACTCTCGCGCTTC 198
QY      1189 CTATCTCGAAGCTTTGTAAGGAGTCTCTGCTGTAGACCCCTAGTCCGTTATGCTA 1248
Db      199  CTTACTTGACACCGTGTGAAGGAGACTTTGCGTGTACACCCCTCCCTTTGGAGAGACCA 258
QY      1249 ACCGTGAATGCTTAAGAGTGAAGACTTTCATCCACCTTTCGCGACCTGTCTATTGGTCGAG 1308
Db      259  TAGCGG---TCGCGATGAAGACGATATTCTGCGCTCGAGAGCCACTCAAGGACAAAC 315
QY      1309 ATGGTTCGGTATCAACAGAGTTCGGATCAGAAAGNACGATGGTATGCTTCGGTTGT 1368
Db      316  ATGGCGTGTACACACGCGAATTAGGATCAGAAAGGACCGCGATCTCTATCCCGATCT 375
QY      1369 TCAACATCAATCGTTCAAAGTTTCATTTATGGAGAAGATGCGAAGAATTCAGACCGGAGA 1428
Db      376  T---GATGAACCGTTCGAAGGGTTGTGGGCCCTGACGCCACGAGTTCAAACCGGAAC 432
QY      1429 GGTGGCTTGAGG 1440
Db      433  GCTGGACGACG 444

RESULT 3
CD275923
LOCUS
DEFINITION
  T143B00226 (FHIG:B) Axenic plate culture Paxillus involutus cDNA
  5', mRNA sequence.
ACCESSION
  CD275923
VERSION
  CD275923.1 GI:34387969
KEYWORDS
  EST
SOURCE
  Paxillus involutus
  Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
  Boletales; Paxillineae; Paxillaceae; Paxillus.
REFERENCE
  1 (bases 1 to 475)
  Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
  Lundberg, J., Uhlen, M. and Tunlid, A.
  Transcriptional responses of Paxillus involutus and Betula pendula
  during formation of ectomycorrhizal root tissue
  Unpublished (2003)
  Contact: Johansson, T.
  Fungal-Host Interaction Group (FHIG)
  Microbial Ecology, Institution of Ecology
  Ecology Building, Lund University, SE-223 62 Lund, Sweden
  Tel: +46 46 222 45 49
  Fax: +46 46 222 41 58
  Email: tomas.johansson@mbioekol.lu.se
PCR PRIMERS
  FORWARD: P104 (5'-GGGAAGCGCGCCATTGTGT-3')
  BACKWARD: P105 (5'-AGTAGCTCGAATCGGCC-3')
Seq primer: P104
High quality sequence stop: 475.
Location/Qualifiers
  1. .475
  /organism="Paxillus involutus"
  /mol_type="mRNA"
  /strain="ATCC 200175"
  /db_xref="taxon:71150"
  /tissue_type="mycelium"
  /dev_stage="25 days of growth after transfer"
  /lab_host="Escherichia coli BM25.8"
  /clone_lib="(FHIG:B) Axenic plate culture"
  /notes="Vector: pTriplex2; Site 1: Sfil; Site 2: Sfil; This
  EST clone is originating from one of three cDNA libraries,
  constructed for transcript profiling of the mycorrhizal
  interaction between the basidiomycete Paxillus involutus
  and Betula pendula (birch). One library represents the
  developed and functional mycorrhizal root tissue

```

## FEATURES

source

```

(' (FHIG:A) Ectomycorrhiza plate culture'), a second
library represents axenically grown fungus (' (FHIG:B)
Axenic plate culture') and a third library represents
axenically grown plants (' (FHIG:C) Axenic plate culture').
Libraries were analyzed in parallel and 3555 (FHIG:A),
3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)
ESTs of >99bp have been deposited. The cDNA libraries were
constructed from total RNA using the SMART cDNA library
construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)
according to the manufacturer's instructions. Full-length
cDNAs were trimmed by Sfil, fractionated and directionally
ligated into (lambda)Triplex2 arms. The lambda library was
converted to a plasmid library via site-specific
recombination at loxP sites in a Cre+ strain (E. coli
BM25.8). Plasmid clones were randomly collected and
analysed by DNA sequencing using a plasmid-specific
forward primer (P104)."

ORIGIN
Query Match      4.1%; Score 79.2; DB 14; Length 475;
Best Local Similarity 54.6%; Pred. No. 0.0025;
Matches 203; Conservative 0; Mismatches 163; Indels 6; Gaps 2;

QY      1069 CAGTCTTGACATGATGTTTACCGGACTCTCAGAAGACAAGCCGTTTCAGGATAAACTTC 1128
Db      88  CAGCGATAAATTTGGCGCTGCTGATGTTGCTAGCCCTGGGATACAGAGAACTCC 147
QY      1129 GAGAAGAAATTTGTCAGATCGACACGGATATGCTCTAGCTAGACGAATTAATGCGTTGC 1188
Db      148  GGGAGGAACCTTCTTCTGTTGACACGGAACGCCCTCGATGGAGAACCTCTCGCGCTTC 207
QY      1189 CTTATCTCGAAGCGTTTGTAAAGAGTCTTTCGTCTAGACCCCTCTAGTCCGTTATGCTA 1248
Db      208  CTTACCTTGACACCGCTGTGTAAGGAGACTTTGCGTGTACACCCCTCCCTTTGGAGAGACA 267
QY      1249 ACCGTGAATGCTTAAAGGATGAAGACTTTCATCCACTTCCGAGCCTGTCTATGGTCGAG 1308
Db      268  TAGCGG---TCGCGATGAAGACGATATTCTGCGCTCGAGAGACCACTCAGGACAAAC 324
QY      1309 ATGGTTCGGTTCATCAACGAGTCCGGATCAGAAAGAACGATGGTTCATGCTCCGTTGT 1368
Db      325  ATGGCGTGTACACGCGGAATTAGGATCAGCAAGGAGCCGCGATCTCTATCCCGATCT 384
QY      1369 TCAACATCAATCGTTCAAAGTTTCATTTATGGAGAAGATGCGAAGAATTCAGACCGGAGA 1428
Db      385  T---GATGAACCGTTCGAAGGGTTGTGGGCCCTGAGCCCGAGTTCAAACCGGAAC 441
QY      1429 GGTGGCTTGAGG 1440
Db      442  GCTGGGACGACG 453

RESULT 4
AT002896
LOCUS
DEFINITION
  AT002896 POSLM01 Pleurotus ostreatus cDNA clone 355LM, mRNA
  sequence.
ACCESSION
  AT002896
VERSION
  AT002896.1 GI:6934623
KEYWORDS
  EST.
SOURCE
  Pleurotus ostreatus
  (oyster mushroom)
  Pleurotus ostreatus
  Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
  Agaricales; Pleurotaceae; Pleurotus.
REFERENCE
  1 (bases 1 to 615)
  Lee, S.H., Kim, B.G., Kim, K.J., Lee, J.S., Yun, D.W., Hahn, J.H.,
  Kim, G.H., Lee, K.H., Suh, D.S., Kwon, S.T., Lee, C.S. and Yoo, I.B.
  Comparative Analysis of Sequences Expressed during the
  Liquid-Cultured Mycelia and Fruit Body Stages of Pleurotus
  ostreatus
  Fungal Genet. Biol. 35 (2), 115-134 (2002)
MEDLINE
  2183865
PUBMED
  11848675

```

COMMENT Contact: Kim BG  
Division of applied microbiology  
National Institute of Agricultural Science and Technology (NIAT)  
Seodundeong, Suwon, Kyung-gi 441-707, South Korea  
Email: bgkimniat.go.kr  
Submitted through BRIC(Biological Research Information Center) of Korea  
URL: http://bric.postech.ac.kr/.  
Location/Qualifiers  
1. .615  
/organism="Pleurotus ostreatus"  
/mol\_type="mRNA"  
/culivar="ASI 2029"  
/db\_xref="taxon:5322"  
/clone="J55LM"  
/dev\_stage="shaking liquid cultured mycelia"  
/lab\_host="E.coli"  
/clone\_lib="POSIM01"  
/note="Vector: lambda Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; average insert size:1000 bp;initial pfu:5 X 10<sup>7</sup>  
Library information:Isolation of total RNA from the mycelia incubated in shaking liquid MCM media at 30 deg C"

FEATURES  
source  
1. .615  
/organism="Pleurotus ostreatus"  
/mol\_type="mRNA"  
/db\_xref="taxon:5322"  
/clone="J55LM"  
/dev\_stage="shaking liquid cultured mycelia"  
/lab\_host="E.coli"  
/clone\_lib="POSIM01"  
/note="Vector: lambda Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; average insert size:1000 bp;initial pfu:5 X 10<sup>7</sup>  
Library information:Isolation of total RNA from the mycelia incubated in shaking liquid MCM media at 30 deg C"

ORIGIN  
Query Match 3.9%; Score 75.8; DB 9; Length 615;  
Best Local Similarity 48.9%; Pred. No. 0.008;  
Matches 234; Conservative 0; Mismatches 242; Indels 3; Gaps 1;  
QY 1265 GGATGAAGACTTCATCCCATTCGCGAGCCTGCTGTTGGTCGAGATGGTGGTCAATCA 1324  
Db 59 GGTGATGATGTTCTCCCGCTTGAGCGTCCGTTGACAGATCTGAAGGGTATTACTCATCA 118  
QY 1325 CGAGTCCGGATCAGGAAGGAACGATGCTGCTCCGTTGTTCAACATCAATCGTTC 1384  
Db 119 ATCAGTAGCAATCAAGAAGACAAATGTTGATGTTCCGATCTTCGGATGATGAT 178  
QY 1385 AAAGTTCAATTAAGAGAGATGCAAGAATTCAGACCGGAGAGTGGCTTGAGGACGT 1444  
Db 179 GGTATCAATATGGGCGAGGATGCTTTTGAATTCAGCCAGACAGCTTGGC---AATCACC 235  
QY 1445 AACAGACTCGCTCAACAGTATTGAACACCTATGGACACGAGCGAGCTTTATCTCTGG 1504  
Db 236 GCGGAGGCTGCACGGTCTCCAGGATATGAGTACATGATGATGATGATGATGATGATG 295  
QY 1505 ACCAGAGCTGCTTTGGTGGCGATTTGCTGTCGCGAGATGAAGCGCTTTGTTGTTG 1564  
Db 296 TCTCGAGCATGATATCGGTTGAGATTTCTATCGTAGAGATGAAGCGCTTGTGTTTCA 355  
QY 1565 CACTCTCGTGGTCCAGTTCGAGCCATCATCTCTCATCCAGATACGAGCATCAC 1624  
Db 356 CTTGTGCGTTCAATCGAGTTCGAGTTCGCTGTTCCATCAGAGAAATTTGGCAAGATCTC 415  
QY 1625 CTTGATCATTTCCGCTCCGTAATCGTTGGTAGAGAGAGAGGGTACCAGATGCGTTT 1684  
Db 416 GTCCATCGTCGAGGACCTATTCTGAAGAACGATAAGAGCGGGAATGTGATGCCCTT 475  
QY 1685 GCAGGTCAAGCGGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1743  
Db 476 GATCATTAACCATATATCAACGGAGTGACCTCAACATTTAACGGATGATATGATGCGG 534

RESULT 5  
CD274181  
LOCUS  
DEFINITION  
Ti43B02839F (FHIG:B) Axenic plate culture Paxillus involutus cDNA  
5', mRNA sequence.  
ACCESSION  
CD274181  
VERSION  
CD274181.1 GI:34386227  
KEYWORDS  
EST.  
SOURCE  
Paxillus involutus  
Paxillus involutus  
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
Boletales; Paxillineae; Paxillaceae; Paxillus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 354)  
Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,  
Lundeberg, J., Uhlen, M. and Tunlid, A.  
Transcriptional responses of Paxillus involutus and Betula pendula  
during formation of ectomycorrhizal root tissue  
Unpublished (2003)  
Contact: Johansson, T.  
Fungal-Host Interaction Group (FHIG)  
Microbial Ecology, Institution of Ecology  
Ecology Building, Lund University, SE-223 62 Lund, Sweden  
Tel: +46 46 222 45 49  
Fax: +46 46 222 41 58  
Email: tomas.johansson@bioekol.lu.se  
PCR Primers  
FORWARD: P104 (5'-GGGAAGCGCGCCATTGTGTT-3')  
BACKWARD: P105 (5'-AGTAGCTCGAATTCGGGCC-3')  
Seq primer: P104  
High quality sequence stop: 354.  
Location/Qualifiers  
1. .354  
/organism="Paxillus involutus"  
/mol\_type="mRNA"  
/strain="ATCC 200175"  
/db\_xref="taxon:71150"  
/tissue\_type="Mycelium"  
/dev\_stage="25 days of growth after transfer"  
/lab\_host="Escherichia coli BM25.8"  
/clone\_lib="(FHIG:B) Axenic plate culture"  
/note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This  
EST clone is originating from one of three cDNA libraries,  
constructed for transcript profiling of the mycorrhizal  
interaction between the basidiomycete Paxillus involutus  
and Betula pendula (birch). One library represents the  
developed and functional mycorrhizal root tissue  
(('FHIG:A) Ectomycorrhiza plate culture'), a second  
library represents axenically grown fungus ('(FHIG:B)  
Axenic plate culture') and a third library represents  
axenically grown plants ('(FHIG:C) Axenic plate culture').  
Libraries were analyzed in parallel and 3555 (FHIG:A),  
3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)  
ESTs of >99bp have been deposited. The cDNA libraries were  
constructed from total RNA using the SMART cDNA library  
construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)  
according to the manufacturer's instructions. Full-length  
cDNAs were trimmed by SfiI, fractionated and directionally  
ligated into (lambda)Triplex2 arms. The lambda library was  
converted to a plasmid library via site-specific  
recombination at loxP sites in a Cre+ strain (E. coli  
BM25.8). Plasmid clones were randomly collected and  
analysed by DNA sequencing using a plasmid-specific  
forward primer (P104)."

ORIGIN  
Query Match 3.6%; Score 69.2; DB 14; Length 354;  
Best Local Similarity 54.4%; Pred. No. 0.12;  
Matches 162; Conservative 0; Mismatches 133; Indels 3; Gaps 1;  
QY 1069 CAGCTTTGACATGATGTTTACCGACTCTCAGAAGACAAGCGTTCAGGATAACTTC 1128  
Db 60 CAGCGATAACTTGGCGCGCTGCATGAGTTGTCGTAGCCCTGGGATACAGGAAGCTCC 119  
QY 1129 GAGAAGAAATTTGTTCAGATCGACATGCGGATATGCTACGTAGACGAACTTAAATCGCTTC 1188  
Db 120 GGGAGGAATTTCTTCTGTTGACACGGAAGCGCCCTCGATGGACGAACTTCGGCGCTTC 179  
QY 1189 CTTATCTCGAAGCGTTTGTTHAAGAGTCTCTTGTCTAGACCCCTCTAGTCCGATGCTTA 1248  
Db 180 CCTACCTTGACACCGCTGGTGAAGGAGACTTGGCGTGTACACCCCTCCCTTTGTAGAGACCA 239  
QY 1249 ACCGTGAATGCTTAAAGGATGAAGACTTCATCCACCTTGGCGAGCTGTCAATGGTCGAG 1308  
Db 240 TACGCG---TCGCGATGAAGACCATATTCTGCGCTCGAGAGGCCACTCACGACAAC 296

QY 1309 ATGGGTGGTTCATCAACGAGTCCGGATCCAGAAAGGAGGATGGTTCGCTT 1366  
 |||||  
 Db 297 ATGGCGTTCGTACACGACGGAATAGGATACGAAAGGACCGCGATCTCTATCCCGAT 354  
 |||||

CD273194 403 bp mRNA linear EST 01-SEP-2003  
 T143B00316 (FHIG:B) Axenic plate culture Paxillus involutus cDNA  
 5', mRNA sequence.

ACCESSION CD273194  
 VERSION CD273194.1 GI:34385240  
 KEYWORDS EST.  
 SOURCE Paxillus involutus

ORGANISM  
 Paxillus involutus  
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 Boletales; Paxillineae; Paxillaceae; Paxillus.

REFERENCE 1 (bases 1 to 403)  
 AUTHORS Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,  
 Lundberg, J., Uhlen, M. and Tunlid, A.  
 TITLE Transcriptional responses of Paxillus involutus and Betula pendula  
 during formation of ectomycorrhizal root tissue

JOURNAL Unpublished (2003)  
 COMMENT Contact: Johansson, T.  
 Fungal-Host Interaction Group (FHIG)  
 Microbial Ecology, Institution of Ecology  
 Ecology Building, Lund University, SE-223 62 Lund, Sweden  
 Tel: +46 46 222 45 49  
 Fax: +46 46 222 41 58  
 Email: tomas.johansson@mbioekol.lu.se

PCR PRIMERS  
 FORWARD: P104 (5'-GGGAGCGCGCCATTGTGT-3')  
 BACKWARD: P105 (5'-AGTGAGCTCGAATGCGGCC-3')  
 Seq primer: P104  
 High quality sequence stop: 403.

FEATURES  
 Location/Qualifiers  
 1..403  
 /organism="Paxillus involutus"  
 /mol\_type="mRNA"  
 /strain="ATCC 200175"  
 /db\_xref="taxon:71150"  
 /tissue\_type="mycelium"  
 /dev\_stage="25 days of growth after transfer"  
 /lab\_host="Escherichia coli BM25.8"  
 /clone\_lib="(FHIG:B) Axenic plate culture"  
 /note="Vector: pTriplex2; Site 1: Sfil; Site 2: Sfil; This  
 EST clone is originating from one of three cDNA libraries,  
 constructed for transcript profiling of the mycorrhizal  
 interaction between the basidiomycete Paxillus involutus  
 and Betula pendula (birch). One library represents the  
 developed and functional mycorrhizal root tissue  
 ('(FHIG:A) Ectomycorrhiza plate culture'), a second  
 library represents axenically grown fungus ('(FHIG:B)  
 axenic plate culture') and a third library represents  
 axenically grown plants ('(FHIG:C) Axenic plate culture').  
 Libraries were analyzed in parallel and 355 (FHIG:A),  
 3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)  
 ESTs of >99bp have been deposited. The cDNA libraries were  
 constructed from total RNA using the SMART cDNA library  
 Construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)  
 according to the manufacturer's instructions. Full-length  
 cDNAs were trimmed by Sfil, fractionated and directionally  
 ligated into (lambda)Triplex2 arms. The lambda library was  
 converted to a plasmid library via site-specific  
 recombination at loxp sites in a Cre+ strain (B. coli  
 BM25.8). Plasmid clones were randomly collected and  
 analysed by DNA sequencing using a plasmid-specific  
 forward primer (P104)."

ORIGIN  
 Query Match 3.5%; Score 68; DB 14; Length 403;  
 Best Local Similarity 53.6%; Pred. No. 0.18;  
 Matches 165; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

QY 1069 CAGTCTTCATCGATGTTTCCACCGACTCTCAGAGAGAAACCCCTTCAGGATAAATTC 1128  
 |||||  
 Db 99 CACGGATAACTTGGCGCTGCATGAGTTTGTCTAGCCCTCGGATACAGAGAGAGTCC 158  
 |||||

QY 1129 GAGAAGAAATTTGTACAGATCGACACGGATATGCCTACGCTAGACGAACCTTAATGCGTTGC 1188  
 |||||  
 Db 159 GGGAGGAATCTTTCTGTGTGACACGAAAGCCCTCGATGACGAATCTCGGGCTTC 218  
 |||||

QY 1189 CTTATCTCGAAGGTTTGTAGGAGTCTTCTGCTAGACCCCTCTAGTCCGTATGCTA 1248  
 |||||  
 Db 219 CCTACCTTGACACCGTGGGAGGAGACATTTGCGTGTACACCTTCCCTTTGGAGAGACCA 278  
 |||||

QY 1249 ACCGTGAATGCTTAAAGGATGAAGACTTCATCCACTTCCGAGGCTGTCTATGTCGAG 1308  
 |||||  
 Db 279 TACGCG---TCGCGATGAAGAGGATATTTGCGGCTCGAGAGCCACTCAGGACAAAC 335  
 |||||

QY 1309 ATGGGTTCGTCATCAACGAGTCCGGATCCGATCAGMAAGGAACGATGTCATGCTTCGTTGT 1368  
 |||||  
 Db 336 ATGGCGTCGTACACGACGGAATTAGGATCAGCAGAGGACCGCGATCTCTATCCCGATCT 395  
 |||||

QY 1369 TCACATC 1376  
 |||||  
 Db 396 TGATGAAC 403  
 |||||

RESULT 7  
 LOCUS BQ102580 487 bp mRNA linear EST 12-APR-2002  
 DEFINITION MIM155 MIN Nitrogen-replete Schizophyllum library Schizophyllum  
 commune cDNA 5' similar to cytochrome p450, mRNA sequence.

ACCESSION BQ102580  
 VERSION BQ102580.1 GI:20144039  
 KEYWORDS EST.  
 SOURCE Schizophyllum commune  
 ORGANISM Schizophyllum commune  
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 Agaricales; Schizophyllaceae; Schizophyllum.

REFERENCE 1 (bases 1 to 487)  
 AUTHORS Jackson, E.N., Abner, J., Honaas, L., Hittinger, C.T., Green, A.,  
 Lilly, W.W. and Gathman, A.C.  
 TITLE Expressed sequence tags from Schizophyllum commune nitrogen-replete  
 and nitrogen-limited libraries, 2002

JOURNAL Unpublished (2002)  
 COMMENT Contact: Gathman AC  
 Biology Department  
 Southeast MO State University  
 1 University Plaza, Cape Girardeau, MO 63701, USA  
 Tel: 5736512361  
 Fax: 5739866433  
 Email: agathman@semo.edu  
 Seq primer: T3  
 POLYA=No.

FEATURES  
 Location/Qualifiers  
 1..487  
 /organism="Schizophyllum commune"  
 /mol\_type="mRNA"  
 /strain="4-40"  
 /db\_xref="taxon:5334"  
 /tissue\_type="mycelium"  
 /clone\_lib="MIN Nitrogen-replete Schizophyllum library"  
 /note="Vector: lambda Zap; Site 1: EcoRI; Site 2: XhoI;  
 4-day-old mycelia of Schizophyllum commune were  
 transferred from minimal (nitrogen-replete) medium to  
 fresh minimal medium. RNA was extracted twelve hours after  
 transfer and cDNAs prepared."

ORIGIN  
 Query Match 3.2%; Score 61; DB 13; Length 487;  
 Best Local Similarity 48.4%; Pred. No. 2.4;  
 Matches 240; Conservative 0; Mismatches 241; Indels 15; Gaps 2;  
 QY 614 GGATTGGTTCGGTCGAGCTACTCTGGAGCTCATGGCTCTTGCAGGATTGACTATAAGAG 673

```
Db 1 GCATTGGATCAGCCGCTGCTACCTTTGATGCTCTCGGTGTCGACGGCTTCGACATACTT 60
QY 674 CGACTCGCTCCAGACACAGACCAATGAGCTCTATGTCGCTTTTGTGCGACTTACCGATGG 733
Db 61 CGAGGCGATCCACACGAGGACATGAGCTGCTCAATGCGTATTAAGACATGTCGAACT 120
QY 734 GTTTGCTCTCACTTGGAGCTGCTTCAAGGCTATCATGTTGGGATTTTGTACTTACTTCCG 793
Db 121 GGTGTGTGTCGAAGCAGCTTCTGCGACAGCATGACAAACATTTACGCGCGTCTTGTCT 180
QY 794 AACTATGAACGGAGACATGAGATAC---CTTTGACTCAAGGATTAGCAGTTTCCCGACG 850
Db 181 AGCCTCTTCCCTACCCAGAACAGCGGACCGTGAAACGCTGCCGAGAGTCATCCGGCG 240
QY 851 AGTGGGATCGAGCTTATGAGCAAAAGAGACAGCGCGTGTGCTTGGCTCAGCTTCCGATCA 910
Db 241 CGTACGCGGCGACATCATCCAGCACACAGAAGC-----GCAAGATCGAGGAGGG 288
QY 911 GCGTGTGTATAAAAGGATGTTCAAGGTCGGGATATCTTAAGTCTCTAGTGAGAGCAA 970
Db 289 CATGCGCGAGCGCAAAACCTTAGCGGCCAAGGACATCTCACCCTCTCTGCTCAAGTCGAA 348
QY 971 CATGCGCGCAACCTGCTGCTGAATCTCAAAAGCTGTCCGATGAGGAGGTACTCGCTCAGT 1030
Db 349 CATGCTCCACGACATCCGCGCCGACAGCGCATCACCGACGCCGACCTCTCTCGACACAT 408
QY 1031 CAGTAACCTGTTATTTGCTGGATATGAACCTTCTCGACAGCTTTCAGATGATGTTTCA 1090
Db 409 CAACAGCTGCGCTTGGCGGGTCCGACACCTCGTCCCTCGCTCACCCTGAGCGCTCCA 468
QY 1091 CCGACTCTCAGAAGAC 1106
Db 469 CTTCTCTCGCGAGCAC 484

RESULT 8
CD275427 400 bp mRNA linear EST 01-SEP-2003
LOCUS T143B04144F (FHIG:B) Axenic plate culture Paxillus involutus cDNA
DEFINITION 5', mRNA sequence.
ACCESSION CD275427
VERSION CD275427.1 GI:34387473
KEYWORDS EST.
SOURCE Paxillus involutus
ORGANISM Paxillus involutus
REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
AUTHORS Boletales; Paxillineae; Paxillaceae; Paxillus.
TITLE 1 (bases 1 to 400)
JOURNAL Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
COMMENT Lundberg, J., Uhlen, M. and Tunlid, A.
Transcriptional responses of Paxillus involutus and Betula pendula
during formation of ectomycorrhizal root tissue
Unpublished (2003)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@bioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAGCGCGCCATGCTGT-3')
BACKWARD: P105 (5'-AGTGAGCTCGAATTCGGCC-3')
Seq primer: P104
High quality sequence stop: 400.
Location/Qualifiers
1. 400
/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 500175"
/db_xref="taxon:71150"
/tissue_type="Mycelium"

FEATURES
source
```

```
/dev_stage="25 days of growth after transfer"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(FHIG:B) Axenic plate culture"
/note=vector: pTriplex2; Site: 1: Sfil; Site: 2: Sfil; This
EST clone is originating from one of three cDNA libraries,
constructed for transcript profiling of the mycorrhizal
interaction between the basidiomycete Paxillus involutus
and Betula pendula (birch). One library represents the
developed and functional mycorrhizal root tissue
('FHIG:A) Ectomycorrhiza plate culture'), a second
library represents axenically grown fungus ('FHIG:B)
Axenic plate culture') and a third library represents
axenically grown plants ('FHIG:C) Axenic plate culture').
Libraries were analyzed in parallel and 3555 (FHIG:A),
3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)
ESTs of >99bp have been deposited. The cDNA libraries were
constructed from total RNA using the SMART cDNA library
Construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)
according to the manufacturer's instructions. Full-length
cDNAs were trimmed by Sfil, fractionated and directionally
ligated into (lambda)Triplex2 arms. The lambda library was
converted to a plasmid library via site-specific
recombination at loxP sites in a Cre+ strain (E. coli
BM25.8). Plasmid clones were randomly collected and
analysed by DNA sequencing using a plasmid-specific
forward primer (P104)."

ORIGIN
Query Match 3.0%; Score 58.8; DB 14; Length 400;
Best Local Similarity 59.6%; Pred. No. 5.9; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 67;

QY 1069 CAGTCTTCACATGATGTTTCACGCACTCTCAGAAGCAAGCGCTTCAGATPAACATTC 1128
Db 52 CAGCGATTAATCTGGCGCTGCTGATGTTGCTAGCCCTCGGATACAGAGGAGCTCC 111
QY 1129 GAGAAGAAATTTGTCAGATCGACCGATATGCTAGCTAGACGACATTAAATGCGTTGC 1188
Db 112 GCGAGGAATCTCTTCTGTTGACACGGAAGCGCCCTCGATGGACGAACTCTCGCGCTTC 171
QY 1189 CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTGCTAGACCCCTCC 1234
Db 172 CTTACTTGCACCGTGGTGAAGGAGACTTGGGTGTACACCCCTCC 217

RESULT 9
CD275197 317 bp mRNA linear EST 01-SEP-2003
LOCUS T143B015624F (FHIG:B) Axenic plate culture Paxillus involutus cDNA
DEFINITION 5', mRNA sequence.
ACCESSION CD275197
VERSION CD275197.1 GI:34387243
KEYWORDS EST.
SOURCE Paxillus involutus
ORGANISM Paxillus involutus
REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
AUTHORS Boletales; Paxillineae; Paxillaceae; Paxillus.
TITLE 1 (bases 1 to 317)
JOURNAL Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
COMMENT Lundberg, J., Uhlen, M. and Tunlid, A.
Transcriptional responses of Paxillus involutus and Betula pendula
during formation of ectomycorrhizal root tissue
Unpublished (2003)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@bioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAGCGCGCCATGCTGT-3')
BACKWARD: P105 (5'-AGTGAGCTCGAATTCGGCC-3')
Seq primer: P104
High quality sequence stop: 400.
Location/Qualifiers
1. 400
/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 500175"
/db_xref="taxon:71150"
/tissue_type="Mycelium"

FEATURES
source
```

Seq primer: P104

High quality sequence stop: 317.

Location/Qualifiers

# FEATURES

source

1. 317  
/organism="Xenopus laevis"

/mol\_type="mRNA"

/strain="ATCC 200175"

/db\_xref="taxon:71150"

/tissue\_type="mycelium"

/dev\_stage="25 days of growth after transfer"

/lab\_host="Escherichia coli BM25.8"

/clone\_lib="(PHIG:B) Axenic plate culture"

/note="(Vector: pTriplex2; Site 1: Sfil; Site 2: Sfil; This

EST clone is originating from one of three cDNA libraries,

constructed for transcript profiling of the mycorrhizal

interaction between the basidiomycete Paxillus involutus

and Betula pendula (birch). One library represents the

developed and functional mycorrhizal root tissue

(' (PHIG:A) Ectomycorrhiza plate culture'), a second

library represents axenically grown fungus (' (PHIG:B)

axenic plate culture') and a third library represents

axenically grown plants (' (PHIG:C) Axenic plate culture').

Libraries were analyzed in parallel and 3555 (PHIG:A),

3964 (PHIG:B), and 2532 (PHIG:C) high-quality (PHRED 20)

ESTs of >99bp have been deposited. The cDNA libraries were

constructed from total RNA using the SMART cDNA library

construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)

according to the manufacturer's instructions. Full-length

cDNAs were trimmed by Sfil, fractionated and directionally

ligated into (lambda)Triplex2 arms. The lambda library was

converted to a plasmid library via site-specific

recombination at loxp sites in a Cre+ strain (E. coli

BM25.8). Plasmid clones were randomly collected and

analysed by DNA sequencing using a plasmid-specific

forward primer (P104)."

## ORIGIN

Query Match 3.0%; Score 58; DB 14; Length 317;

Best Local Similarity 55.4%; Pred. No. 8.7;

Matches 112; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 1069 CAGTCTTGACATGGATGTTTCCAGCACTCTCAGAGACAAAGCCGCTTCAGGATAAATTC 1128

Db 60 CACGGATACTGGGCGCTGCATGAGTTGTCCTAGCCCTGGGATACAGAGAGCTCC 119

QY 1129 GAGAGAAATTTGTTCAGATCGACAGATATGCCATCGCTAGACGAACTTAATCGTTGC 1188

Db 120 GGAGGAACTCTTCTGTTGACACGAAACGCCCTCGATGGACGAACTCTCGCGCTTC 179

QY 1189 CTTATCTCGAGGTTTGTAGAGGATCTCTTCGCTAGACCTCTAGTCCGTATGCTA 1248

Db 180 CTTACCTTGACACCGTGGTGAAGGAGACTTTGCGTGTGACCCCTCTTGAGAGACCA 239

QY 1249 ACCGTGAATCTTTAAAGGATGA 1270

Db 240 TACGCGTCGGATGAAGACGA 261

## RESULT 10

BJ075377

LOCUS

DEFINITION BJ075377 NIBB Mochii normalized Xenopus tailbud library Xenopus

laevis cDNA clone XL057f08 5', mRNA sequence.

ACCSSION BJ075377.1 GI:17520293

VERSION BJ075377.1

KEYWORDS EST.

SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 499).

REFERENCE Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and

Kohara,Y.

Expressed genes in X. laevis embryo

Unpublished (2001)

Contact: Tadashi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp

The information of this clone is available through the following

URL.

http://xenopus.nibb.ac.jp.

Location/Qualifiers

1. 499

/organism="Xenopus laevis"

/mol\_type="mRNA"

/db\_xref="taxon:8355"

/clone\_lib="XLO57f08"

/tissue\_type="whole embryo"

/dev\_stage="stage 25"

/clone\_lib="NIBB Mochii normalized Xenopus tailbud

library"

ORIGIN

Query Match 2.9%; Score 56.8; DB 12; Length 499;

Best Local Similarity 53.8%; Pred. No. 11;

Matches 143; Conservative 0; Mismatches 117; Indels 6; Gaps 1;

QY 996 CAAAAGCTGTCGGATGAGGAGTACTCGCTCAGATCACTGTAACCTGTTATTGCTGGATAT 1055

Db 52 CAGAGCTTTCTGATGAGGATATACGGCTCAGTGCGACACATTCATGTTTGAGGACAT 111

QY 1056 GAAATCTTTCGACAGCTTTCAGATGGATGTTTCCAGCATCTCAGAGACAAGCCGTT 1115

Db 112 GATACATCATCTAGTGAATATCTCTGGATCTCTACATTTGTATGCTATCCAGACAC 171

QY 1116 CAGGATAAATTCGAGAGAAATTTGTCAGATCGACACGATATGCTAGCTAGAC--- 1172

Db 172 CAGCAGAAATGCGAGGAGAGATCAATGAGTCTTGGGGAGCGACAACCATGGAATGG 231

QY 1173 ---GAACCTTAATGCGTTCCCTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGAC 1229

Db 232 GATGACCTTAATAGGATACCTTATACCAATGTGTATTAAAGAAAGTCTCGCGCTATAC 291

QY 1230 CCTCCTAGTCCGTATGCTAACCGTGA 1255

Db 292 CCTCCAGTCCCATCATGATCCAGGA 317

## RESULT 11

BJ079027/c

LOCUS

DEFINITION

EX779027 XGC-egg Silurana tropicalis cDNA clone TEGG070n12 3', mRNA

sequence.

ACCSSION EX779027

VERSION EX779027.1

KEYWORDS GI:39686233

SOURCE EST.

ORGANISM

Silurana tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 971)

REFERENCE Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11\_2003)

UNPUBLISHED (2003)

CONTACT: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS\_SEQUENCE\_ID: TEGG070n12.q1kt7

Sequencing primer: T7  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn  
cDNA was oligo dT primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the  
5' end and NotI at the 3' end.  
Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
Host: Escherichia coli XL1-blue.  
Location/Qualifiers

## FEATURES

## source

1. .971  
/organism="Silurana tropicalis"  
/mol\_type="rRNA"  
/db\_xref="taxon:8364"  
/clone="TEGG070n12"  
/dev\_stage="egg"  
/lab\_host="Escherichia coli XL1-blue"  
/clone\_lib="XGC-egg"  
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
was oligo dT primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with  
EcoRI at the 5' end and NotI at the 3' end"

## ORIGIN

Query Match 2.9%; Score 55.4; DB 13; Length 971;  
Best Local Similarity 54.7%; Pred. No. 15;  
Matches 135; Conservative 0; Mismatches 106; Indels 6; Gaps 1;  
QY 995 TCAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATA 1054  
Db 817 TAAAGCTCTGACAGATGCAGAAATAATGCTCAGTCCCTTATATTTATATTCGCTGGATA 758  
QY 1055 TGAACCTCTTCGACAGTCTTACATGGATGTTTCACCGACTCTCAGAAGACAAAGCCGT 1114  
Db 757 TGAACCAACAGTACAGACTCTTACTGCTTTTATATCTTGCACCCACCTGACGT 698  
QY 1115 TCAGGATAAAGCTCGAAGAAATTT-----TGTGAGATCGACAGGATATGCTTACGCT 1168  
Db 697 CCAGCAGAAATCGCATGAGGATAGATTCATTTCTCCCTGACAGGCCAGCTCTACTTA 638  
QY 1169 AGACGAACCTTAATGGTTCCTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGA 1228  
Db 637 TGACATTTAATGCAGATGGAATATCTTGATATGGTTATTCAGGAGACCTCAGGTGTT 578  
QY 1229 CCTCTCT 1235  
Db 577 TCCCTCT 571

## RESULT 12

CF152584  
LOCUS  
DEFINITION  
AGENCOURT.14907205 NICHD XGC Emb8 Silurana tropicalis cDNA clone  
IMAGE:6985215 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Silurana tropicalis (western clawed frog)  
Silurana tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Silurana.

## REFERENCE

1. (bases 1 to 849)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: Robert M. Grainger  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LMA14652 row: h column: 14  
High quality sequence start: 66  
High quality sequence stop: 708.  
Location/Qualifiers

## FEATURES

## source

1. .849  
/organism="Silurana tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="IMAGE:6985215"  
/tissue\_type="tadpole"  
/dev\_stage="embryo, stages 40-45"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHD XGC Emb8"  
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;  
Cloned unidirectionally. Primer: Oligo dT. Average insert  
size 2.1 kb. Constructed by invitrogen. Note: This is a  
Xenopus Gene Collection (XGC) library."

## ORIGIN

Query Match 2.8%; Score 53.8; DB 14; Length 849;  
Best Local Similarity 54.3%; Pred. No. 29;  
Matches 134; Conservative 0; Mismatches 107; Indels 6; Gaps 1;  
QY 995 TCAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATA 1054  
Db 480 TAAAGCTCTGACAGATGCAGAAATAATGCTCAGTCCCTTATATTTATATTCGCTGGATA 539  
QY 1055 TGAACCTCTTCGACAGTCTTACATGGATGTTTCACCGACTCTCAGAAGACAAAGCCGT 1114  
Db 540 TGAACCAACAGTACAGACTCTTACTGCTTTTATATCTTGCACCCACCTGACGT 599  
QY 1115 TCAGGATAAAGCTCGAAGAAATTT-----TGTGAGATCGACAGGATATGCTTACGCT 1168  
Db 600 CCAGCAGAGATGCTGATGAGGATAGATTCATTTCTCCCTGACAGGCCAGCTCTACTTA 659  
QY 1169 AGACGAACCTTAATGGTTCCTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGA 1228  
Db 660 TGACATTTAATGCAGATGGAATATCTTGATATGGTTATTCAGGAGACCTCAGGTGTT 719  
QY 1229 CCTCTCT 1235  
Db 720 TCCCTCT 726

## RESULT 13

AM600900  
LOCUS  
DEFINITION  
7C2 cDNA library of 4-day-old Eucalyptus globulus  
bicostata-Pisolithus tinctorius ectomycorrhiza Pisolithus  
tinctorius cDNA 5' similar to hydrophobin, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Pisolithus tinctorius  
Pisolithus tinctorius  
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
Boletales; Sclerodermatales; Pisolithaceae; Pisolithus.

## REFERENCE

1. (bases 1 to 568)  
Volbler, C., Duplessis, S., Encelot, N. and Martin, F.  
Identification of symbiosis-regulated genes in Eucalyptus  
globulus-Pisolithus tinctorius ectomycorrhiza by differential  
hybridization of arrayed cDNAs  
Plant J. 25 (2), 181-191 (2001)  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: Martin FM  
Equipe de Microbiologie Forestiere  
Institut National de la Recherche Agronomique  
Centre INRA de Nancy, 54280 Champenoux, France





in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATTTAGCC-3' and 3' adaptor sequence: 5'-ATCTTAGCGGCGGCGACATCT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.7 kb (range 0.8-3.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

Query Match	2.6%; Score 50.4; DB 14; Length 875;
Best Local Similarity	52.3%; Prid. No. 1e+02;
Matches 139; Conservative 0;	Mismatches 121; Indels 6; Gaps 1;
QY	996 CAAAGCTGTCCGATGAGGAGTACTCGCTCAGATCAGTAACTGTATTTCCTGGATAT 1055
Db	101 CAAAGGCTTTCTGATGAGGATATACGCGCTGAGGTGGACACATTCATGTTTGGGGACAT 160
QY	1056 GAAATCTCTTCGACAGTCTTGACATGGATGTTTCACCGACTCTCAGAGACAAAGCCGTT 1115
Db	161 GATACATCATCTAGTGAATATCTGTGATCTTATTTGTATGGCTACATCTCAGAGCAC 220
QY	1116 CAGGATAAATTCGAGAAGAAATTTGTTCAGATGCACACGGATATGCTTACGCTAGAC --- 1172
Db	221 CAGCAGAAATGCCAGGAGGAGATCAACGAAGTCTTGGGGAGCGCAAAACCATGGATGG 280
QY	1173 --- GAACATTATGGTTGCGCTTATTCGAGAGGTTTGTTAAGGAGTCTCTTCGCTAGAC 1229
Db	281 GATGACCTTGGTAGATACCTTATATCCACATGTTGTTAAGGAAAGTCTCGGCTATAC 340
QY	1230 CCTCTAGTCCGTATGCTAAACCGTGA 1255
Db	341 CCTCCAGTCCCATCAGTATCCAGGA 366

Search completed: April 6, 2004, 17:46:05  
Job time : 5202 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 08:11:57 ; Search time 7663 Seconds  
(without alignments)  
10927.671 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932

Sequence: 1 gaattcgacagagccacc.....agccggctcgtcggaattc 1932

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pin.\*

35: em.htg.rod.\*

36: em.htg.man.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1932	100.0	1932	6	AR203449 Sequence
2	1932	100.0	1932	6	AX034665 Sequence
3	353.2	18.3	3969	6	AR203450 Sequence
4	353.2	18.3	3969	6	AX034666 Sequence
5	55	2.8	2514	5	BC060496 Xenopus l
6	51.8	2.7	2000	6	AX655393 Sequence
7	48.4	2.5	1062	9	AX563377 Homo sapi
8	48.4	2.5	1365	9	AX563376 Homo sapi
9	48.4	2.5	1512	6	E10636 Human cDNA
10	48.4	2.5	1512	6	E10855 cDNA encodi
11	48.4	2.5	1512	6	AR399323 Sequence
12	48.4	2.5	1512	9	AX563375 Homo sapi
13	48.4	2.5	2011	6	AX408939 Sequence
14	48.4	2.5	2011	9	HUMCYPHLP
15	48.4	2.5	2011	9	HUMPA450M
16	48.4	2.5	2036	9	HUMCIPNO
17	48.4	2.5	2059	6	AR380517
18	48.4	2.5	2059	9	HUMCYPNOA
19	48.4	2.5	2059	11	G18329
20	48.4	2.5	2759	6	AR222892 Sequence
21	48.4	2.5	2759	9	HSRCYP3
22	48.4	2.5	2849	6	AX409463
23	48.4	2.5	2849	9	HUMCITLIIA
24	48	2.5	154452	2	AC137912
25	47.4	2.5	1410	8	AB084385
26	47.4	2.5	1684	10	BC040779
27	47.4	2.5	1871	8	AK072295
28	47.4	2.5	1900	8	AK101670
29	47.4	2.5	2160	6	AR403280
30	47.4	2.5	2160	6	BD140625
31	47.2	2.4	3588	9	HSM802481
32	46.8	2.4	1512	9	AY334551
33	46.8	2.4	1704	9	AF182273
34	46.8	2.4	1712	6	AR352699
35	46.8	2.4	1712	6	BD195644
36	46.8	2.4	1822	6	AR352698
37	46.8	2.4	1822	6	BD195643
38	46.8	2.4	2032	9	SS3047
39	46.8	2.4	2139	9	BC015814
40	46.8	2.4	6004	10	AF372834
41	46.6	2.4	1395	5	AB091340
42	46.4	2.4	1760	3	AY051533
43	46.2	2.4	1286	9	AF131809
44	46.2	2.4	179437	5	AL928892
45	46.2	2.4	197330	2	AL928858

## ALIGNMENTS

RESULT 1

AR203449

LOCUS

DEFINITION

AR203449

ACCESSION

AR203449

VERSION

AR203449.1

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 1932)

AUTHORS

Hoshino, T., Ojima, K. and Setoguchi, Y.

TITLE

Ataxanthin synthase

JOURNAL

Patent: US 6365386-A 2 02-APR-2002;

Location/Qualifiers

AR203449 1932 bp DNA linear PAT 20-JUN-2002  
Sequence 2 from patent US 6365386.

AR203449.1 GI:21499840

source	1. .1932							
	/organism="unknown"							
	/mol_type="unassigned DNA"							
ORIGIN								
Query Match	100.0%;	Score 1932;	DB 6;	Length 1932;				
Best Local Similarity	100.0%;	Pred. No. 0;						
Matches 1932; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1	GAATTCGGACAGGCGCACCTACTTCTCCATATGTTTCATCTTTGGTCTTGTCTACAGGTG	60					
Db	1	GAATTCGGACAGGCGCACCTACTTCTCCATATGTTTCATCTTTGGTCTTGTCTACAGGTG	60					
QY	61	CTTTAGGCGTGGCTGTTTCTCATGGGATCCATAGCGTTCTTTCAGTCTTTTACCTCGCTC	120					
Db	61	CTTTAGGCGTGGCTGTTTCTCATGGGATCCATAGCGTTCTTTCAGTCTTTTACCTCGCTC	120					
QY	121	CGAGGCGATCTTCACTGTATTAACCTTCAGGCGCCGGAATCATCAACTACTTTACAGGCA	180					
Db	121	CGAGGCGATCTTCACTGTATTAACCTTCAGGCGCCGGAATCATCAACTACTTTACAGGCA	180					
QY	181	ATTTTTTACACATCTCTCAGCTGTACAGGTGAAGAGCATGCGAAGTACAGAGAAAAT	240					
Db	181	ATTTTTTACACATCTCTCAGCTGTACAGGTGAAGAGCATGCGAAGTACAGAGAAAAT	240					
QY	241	ACGGAAGCACCTCCGGTTTGTGGGATCGCTGGAGCACCCGCTCTTGAACCTCGACCGATC	300					
Db	241	ACGGAAGCACCTCCGGTTTGTGGGATCGCTGGAGCACCCGCTCTTGAACCTCGACCGATC	300					
QY	301	CGAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCCGAACCTGGTATGGCGG	360					
Db	301	CGAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCCGAACCTGGTATGGCGG	360					
QY	361	CTCGAGTGTCTCAGAAATTGCTACCGGAGATGGTGTGTTTACGGCGGAGGTGAAGTCTATA	420					
Db	361	CTCGAGTGTCTCAGAAATTGCTACCGGAGATGGTGTGTTTACGGCGGAGGTGAAGTCTATA	420					
QY	421	AGCGACATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGTC	480					
Db	421	AGCGACATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGTC	480					
QY	481	CAATTTCTTAGAAAAGGTATGGAATTTGTCGACCAAGATGATGGAGGATCGGCTCAGA	540					
Db	481	CAATTTCTTAGAAAAGGTATGGAATTTGTCGACCAAGATGATGGAGGATCGGCTCAGA	540					
QY	541	AGGATATGCGCGTGGAGAGTCCGCGGTGAAAAGAGGCAACCAAGACTCGAGACCCGAG	600					
Db	541	AGGATATGCGCGTGGAGAGTCCGCGGTGAAAAGAGGCAACCAAGACTCGAGACCCGAG	600					
QY	601	GAGTCGATGTAAGGATTCGGTCCGTCGAGCTACTCTGGACGTCATGGCTCTTCAGGAT	660					
Db	601	GAGTCGATGTAAGGATTCGGTCCGTCGAGCTACTCTGGACGTCATGGCTCTTCAGGAT	660					
QY	661	TTGACTATTAAGAGCGACTCGCTCCAGAACCAAGCAATGAGCTCTATGTCTTTTGTGCG	720					
Db	661	TTGACTATTAAGAGCGACTCGCTCCAGAACCAAGCAATGAGCTCTATGTCTTTTGTGCG	720					
QY	721	GACTTACCGATGGGTTTGTCTCTACCTTGGACTCGTTCAAGGCTTATCATGTGGGATTTTG	780					
Db	721	GACTTACCGATGGGTTTGTCTCTACCTTGGACTCGTTCAAGGCTTATCATGTGGGATTTTG	780					
QY	781	TACCTTACTTCCGAATATGAACCGGACATGAGATACCTTTGACTCAAGGATTTAGCAG	840					
Db	781	TACCTTACTTCCGAATATGAACCGGACATGAGATACCTTTGACTCAAGGATTTAGCAG	840					
QY	841	TTTCCGACGAGTTGGGATTCGAGCTTATGGAGCAAAAAGAGCGCGCTCTTGGCTCAG	900					
Db	841	TTTCCGACGAGTTGGGATTCGAGCTTATGGAGCAAAAAGAGCGCGCTCTTGGCTCAG	900					
QY	901	CTTCCGATCAGGCTGTTGATAAAAAGGATGTTTCAAGGTCGGGATATCTTAAGTCTCTTAG	960					
Db	901	CTTCCGATCAGGCTGTTGATAAAAAGGATGTTTCAAGGTCGGGATATCTTAAGTCTCTTAG	960					

QY	961	TCAGAGCAACATCGCGCCAACTGCTCAATCTCAAAAGCTGTCCTGATGAGGAGTAC	1020
DB	961	TCAGAGCAACATCGCGCCAACTGCTCAATCTCAAAAGCTGTCCTGATGAGGAGTAC	1020
QY	1021	TCGCTCAGATCAGTAACCTGTTATTTGCTGGATATGAAACTTCTTCGACAGTCTTGACAT	1080
DB	1021	TCGCTCAGATCAGTAACCTGTTATTTGCTGGATATGAAACTTCTTCGACAGTCTTGACAT	1080
QY	1081	GGATGTTTCCAGCTCTCAGAGACAAAGCCGTTCAGGATAAACTTCGAGAGAAATTT	1140
DB	1081	GGATGTTTCCAGCTCTCAGAGACAAAGCCGTTCAGGATAAACTTCGAGAGAAATTT	1140
QY	1141	GTCAGATCGACACCGATATGCTACGTAGACGAACTTAATGCGTTGCTTATCTCGAAG	1200
DB	1141	GTCAGATCGACACCGATATGCTACGTAGACGAACTTAATGCGTTGCTTATCTCGAAG	1200
QY	1201	CGTTGTTAAGGAGTCTCTTCGCTAGACCTCTAGTCCGTAATGCTAACCGTGAATGCT	1260
DB	1201	CGTTGTTAAGGAGTCTCTTCGCTAGACCTCTAGTCCGTAATGCTAACCGTGAATGCT	1260
QY	1261	TAAAGGATGAAGACTTCATCCCACTTCGCGAGCTGTCTATTGGTCGAGATGGTCCGTCA	1320
DB	1261	TAAAGGATGAAGACTTCATCCCACTTCGCGAGCTGTCTATTGGTCGAGATGGTCCGTCA	1320
QY	1321	TCAACGAGGTCGCGATCAAGAAAGAACGATGCTTCCTCGTTGTTCAACATCAATC	1380
DB	1321	TCAACGAGGTCGCGATCAAGAAAGAACGATGCTTCCTCGTTGTTCAACATCAATC	1380
QY	1381	GTTCAAGTTCATTTATGGAGAGATCGAAGAAATTCAGACCGGAGAGTGGCTTGAGG	1440
DB	1381	GTTCAAGTTCATTTATGGAGAGATCGAAGAAATTCAGACCGGAGAGTGGCTTGAGG	1440
QY	1441	ACGTAACAGACTCCCTCAACAGTATTGAAGCACCTATGGACACCCAGGCGAGCTTTATCT	1500
DB	1441	ACGTAACAGACTCCCTCAACAGTATTGAAGCACCTATGGACACCCAGGCGAGCTTTATCT	1500
QY	1501	CTGACCCAGAGCTTCTTGGTGGGATTTGCTGTCGCGAGATGAAGCCCTCTTCT	1560
DB	1501	CTGACCCAGAGCTTCTTGGTGGGATTTGCTGTCGCGAGATGAAGCCCTCTTCT	1560
QY	1561	TTTGACACTCTCCGTCGGGTCCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACA	1620
DB	1561	TTTGACACTCTCCGTCGGGTCCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACA	1620
QY	1621	TCACCTTGATCATTTCCGCTCCTGAAATCGTTGGTAGAGAGAGGAGGTACCAATGC	1680
DB	1621	TCACCTTGATCATTTCCGCTCCTGAAATCGTTGGTAGAGAGAGGAGGTACCAATGC	1680
QY	1681	GTTTGCAGGTCGAAGCGGTGCAATGAGTTCATATGTTTAAAGAGAGTTCATAT	1740
DB	1681	GTTTGCAGGTCGAAGCGGTGCAATGAGTTCATATGTTTAAAGAGAGTTCATAT	1740
QY	1741	CTGAGATGTTGATAGGACAAATGCTTCTTTGATCGATTGTTTTCATACCCGGGC	1800
DB	1741	CTGAGATGTTGATAGGACAAATGCTTCTTTGATCGATTGTTTTCATACCCGGGC	1800
QY	1801	AGGCGCTATGACTTCTAGCTGCTATCGTCGCTGAGCTCTCTTCTTACCCCTATAT	1860
DB	1801	AGGCGCTATGACTTCTAGCTGCTATCGTCGCTGAGCTCTCTTCTTACCCCTATAT	1860
QY	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCGGCT	1920
DB	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCGGCT	1920
QY	1921	CGTCCGGAATTC 1932	
DB	1921	CGTCCGGAATTC 1932	

RESULT 2  
 AX034665  
 LOCUS  
 DEFINITION Sequence 2 from Patent EP1035206.  
 AX034665 1932 bp DNA linear PAT 15-NOV-2000

ACCESSION AX034665  
VERSION AX034665.1 GI:11190638  
KEYWORDS Xanthophyllomyces dendrorhous (anamorph: Phaffia rhodozyma)  
SOURCE Xanthophyllomyces dendrorhous  
ORGANISM Eukaryota; Fungi; Basidiomycota; Hymenomycetes;  
Heterobasidiomycetes; Tremellomycetidae; Cystofilobasidiales;  
Cystofilobasidiaceae; Xanthophyllomyces.  
REFERENCE 1.  
AUTHORS Hoshino, T., Ojima, K. and Setoguchi, Y.  
TITLE Antaxanthin synthetase  
JOURNAL Patent: EP 1035206-A 2 13-SEP-2000;  
HOFFMANN LA ROCHE (CH)  
FEATURES  
Location/Qualifiers  
1..1932  
/organism="Xanthophyllomyces dendrorhous"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5421"  
14..1891  
33..1706  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC16261.1"  
/db\_xref="GI:11190639"  
/db\_xref="REVIEW:161616"  
/translation="MFILVLLTGALGLAFAFWSNIAFPLSLYLAPRRSLVNLQPNHT  
NYFTGNFLDILSARTGEBEHAKYREKYGSTLRFAGIAGAPVLNSDTPKVFHVMKEAYD  
YKPGMAARVLRIATGDGVTVAEGAEHRRHRIIMPSLSAQAVKSVPII FLEKGMELV  
DKMDEADKAVGESAGEKATLEETGVDVDMVGRATLDVMAALGDFYKSDSLQ  
KTNELVAVFVLTDGPAPLIDSPKAIIMWDFVPMKRHEIPLTQGLAVRRVGI  
ELMEKQAVLSASDQAVDKDVQGRDILSLVLANIAPESQKLSDEEVLQIS  
NLIFAGYTSSTVLTWFMHRLSEDRVQDKREBELCQIDMDPLDELNALPYLEAFV  
KESLRDPPSPVANREKDEDFIPLAEVIGRDSVINEVRIKGTWMLPFLNIR  
SKFTYGEDAEERPERWLEVDLSNIEAPYQASFIQSPRACFGWRFAVEMKAF  
LFVTLRRVQPEIISHPEVEHITLIISPRIVGREKEGYMRLOVKPE"  
1871  
polya\_site  
ORIGIN  
Query Match 100.0%; Score 1932; DB 6; Length 1932;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCGGACGAGGCGACCTACTTTCATATGTTTCATCTTGTCTTGTCTTCACAGTG 60  
DB 1 GAATTCGGACGAGGCGACCTACTTTCATATGTTTCATCTTGTCTTGTCTTCACAGTG 60  
QY 61 CTTTAGCGCTGGCTGCTTTCTCATGGCATCCATAGCGTCTTTCAGTCTTACCTCGCTC 120  
DB 61 CTTTAGCGCTGGCTGCTTTCTCATGGCATCCATAGCGTCTTTCAGTCTTACCTCGCTC 120  
QY 121 CGAGGCGATCTTCACTGTATACCTTCAGGCGCCGAATACCACTACTTTACAGGCA 180  
DB 121 CGAGGCGATCTTCACTGTATACCTTCAGGCGCCGAATACCACTACTTTACAGGCA 180  
QY 181 ATTTTATAGATCATCTCTCAGCTCGTACAGGTGAAGAGCATCGAAGTACAGAGAAAT 240  
DB 181 ATTTTATAGATCATCTCTCAGCTCGTACAGGTGAAGAGCATCGAAGTACAGAGAAAT 240  
QY 241 ACCGAGCACCTCCGGTTTGTGGATCGCTGGAGCACCGCTTTGAACCTGACCCGATC 300  
DB 241 ACCGAGCACCTCCGGTTTGTGGATCGCTGGAGCACCGCTTTGAACCTGACCCGATC 300  
QY 301 CGAAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCCGAAACCTGGTATGGCG 360  
DB 301 CGAAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCCGAAACCTGGTATGGCG 360  
QY 361 CTGAGTGCTCAGAAATGCTACCGAGATGTTGTTTACGCGGGAAGTGAGCTCAT 420  
DB 361 CTGAGTGCTCAGAAATGCTACCGAGATGTTGTTTACGCGGGAAGTGAGCTCAT 420  
QY 421 AGCGACATCGAAGATCATGATCCCTCTCTGTGCGCTCAGCGCTTAAGTCGATGTC 480  
DB 421 AGCGACATCGAAGATCATGATCCCTCTCTGTGCGCTCAGCGCTTAAGTCGATGTC 480

QY 481 CAATTTTCTTAAAAAGGCTATGAACTTGTGCAAGATGATCGAGGATCGCGCTCAGA 540  
DB 481 CAATTTTCTTAAAAAGGCTATGAACTTGTGCAAGATGATCGAGGATCGCGCTCAGA 540  
QY 541 AGGATATGCGCGTGGGAGAGTGGCCCGTGAAGAAGGCAACAGACTCGAGACCGAAG 600  
DB 541 AGGATATGCGCGTGGGAGAGTGGCCCGTGAAGAAGGCAACAGACTCGAGACCGAAG 600  
QY 601 GAGTCGATGTAAGGATTTGGGTGCGTGGTCTACTCTGAGACGTCATGGCTCTTTCAGGAT 660  
DB 601 GAGTCGATGTAAGGATTTGGGTGCGTGGTCTACTCTGAGACGTCATGGCTCTTTCAGGAT 660  
QY 661 TTGACTATPAGAGCGACTCGCTCCAGAACAGCAACCAATAGCTCTATGTCTCTTTGTCTG 720  
DB 661 TTGACTATPAGAGCGACTCGCTCCAGAACAGCAACCAATAGCTCTATGTCTCTTTGTCTG 720  
QY 721 GACTTACCGATGGGTGCTCTTACCTTGGCTGGTCTCAAGGCTATCATCTGGGATTTTG 780  
DB 721 GACTTACCGATGGGTGCTCTTACCTTGGCTGGTCTCAAGGCTATCATCTGGGATTTTG 780  
QY 781 TACCTTACTTCCGAACATATGAACCGAGACATGATGATACCTTTGACTCAAGGATTAGCAG 840  
DB 781 TACCTTACTTCCGAACATATGAACCGAGACATGATGATACCTTTGACTCAAGGATTAGCAG 840  
QY 841 TTTCCGACGAGTTGGGATCGAGCTTATGAGCAAAAGAGCAGGCGCTGCTTGGCTCAG 900  
DB 841 TTTCCGACGAGTTGGGATCGAGCTTATGAGCAAAAGAGCAGGCGCTGCTTGGCTCAG 900  
QY 901 CTTCCGATCAGGCTGTTTCATAAAAAGGATGTTCAAGTCCGGGATATCCTTAAGTCTCCTAG 960  
DB 901 CTTCCGATCAGGCTGTTTCATAAAAAGGATGTTCAAGTCCGGGATATCCTTAAGTCTCCTAG 960  
QY 961 TGAGAGCAAAACATCGCCGCCAACCTGCTGAATCTCAAAAGCTGTCGATGAGGAGTAC 1020  
DB 961 TGAGAGCAAAACATCGCCGCCAACCTGCTGAATCTCAAAAGCTGTCGATGAGGAGTAC 1020  
QY 1021 TCGCTCAGATCAGTAAACCTGTTTATTTGCTGGATGAAACTTCTTCGACAGTCTTGACAT 1080  
DB 1021 TCGCTCAGATCAGTAAACCTGTTTATTTGCTGGATGAAACTTCTTCGACAGTCTTGACAT 1080  
QY 1081 GGATGTTTACCGACTCTCAGAGCAAAAGCGCTTCAGGATAAACTTCGAGAGAAATTT 1140  
DB 1081 GGATGTTTACCGACTCTCAGAGCAAAAGCGCTTCAGGATAAACTTCGAGAGAAATTT 1140  
QY 1141 GTCAGATCGACACGATATGCTACGCTAGACGAATTAATGCGTTCGCTTATCTCGAAG 1200  
DB 1141 GTCAGATCGACACGATATGCTACGCTAGACGAATTAATGCGTTCGCTTATCTCGAAG 1200  
QY 1201 CGTTTGTAAAGGAGTCTCTTCTGCTAGACCTCCTAGTCCGATGCTAAACCGTGAATGCT 1260  
DB 1201 CGTTTGTAAAGGAGTCTCTTCTGCTAGACCTCCTAGTCCGATGCTAAACCGTGAATGCT 1260  
QY 1261 TAAAGATGAAGACTTTCATCCACCTTCCGAGCTGCTATGTTGTCGAGATGGGTGCGTCA 1320  
DB 1261 TAAAGATGAAGACTTTCATCCACCTTCCGAGCTGCTATGTTGTCGAGATGGGTGCGTCA 1320  
QY 1321 TCAACGAGTCCGAGTACGAAAGCAATGCTGCTTCCGTTGTTCAACATCAATC 1380  
DB 1321 TCAACGAGTCCGAGTACGAAAGCAATGCTGCTTCCGTTGTTCAACATCAATC 1380  
QY 1381 GTTCAAGTTCATTTATGGAGAAGATGCAAGAATTCAGACCGGAGAGTGGCTTAGG 1440  
DB 1381 GTTCAAGTTCATTTATGGAGAAGATGCAAGAATTCAGACCGGAGAGTGGCTTAGG 1440  
QY 1441 AGCTAACAGCTCGCTCAACAGTATGAGCAACCTATGGACACCGAGCTTTTATCT 1500  
DB 1441 AGCTAACAGCTCGCTCAACAGTATGAGCAACCTATGGACACCGAGCTTTTATCT 1500  
QY 1501 CTGGACCCAGAGCTTGTGTTGGGCTTGTGTCGCGAGATGAAGCCCTTTTGT 1560  
DB 1501 CTGGACCCAGAGCTTGTGTTGGGCTTGTGTCGCGAGATGAAGCCCTTTTGT 1560

	Qy	1561	TTGTGCACTCTCGTTCGGGTCCAGTTCCAGGCCCATCATTCTCTCATCCAGTAGTACGAGCAC	1620
	Dd	1561	TTGTGCACCTCTCGGTTCGGGTCCAGTTCCAGGCCCATCATTCTCTCATCCAGTAGTACGAGCAC	1620
	Qy	1621	TCACTTGTGATCATTTTCGGTCTCTCGAATCGTTGGTAGAGAAGAGGGGTACCAGATGC	1680
	Dd	1621	TCACTTGTGATCATTTTCGGTCTCTCGAATCGTTGGTAGAGAAGAGGGGTACCAGATGC	1680
	Qy	1681	GTTTGCAAGTCAAGCCGGTCCAATGAGTTGAATTCCTCATATGTTAAGAAAGATTCATAT	1740
	Dd	1681	GTTTGCAAGTCAAGCCGGTCCAATGAGTTGAATTCCTCATATGTTAAGAAAGATTCATAT	1740
	Qy	1741	CTGAGAACTGTGACTAGGACAAATGCCTCTCTTGTATCGATTGTGTTCTCTCATACCCGGC	1800
	Dd	1741	CTGAGAACTGTGACTAGGACAAATGCCTCTCTTGTATCGATTGTGTTCTCTCATACCCGGC	1800
	Qy	1801	AGCGCGTATGACTTCTACGTCGTATCGTCGCTCTGGACTCTCTTCTTACCCATATAT	1860
	Dd	1801	AGCGCGTATGACTTCTACGTCGTCTATCGTCGCTCTGGACTCTCTTCTTACCCATATAT	1860
	Qy	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCCGCT	1920
	Dd	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCCGCT	1920
	Qy	1921	CGTGCCGAATTC 1932	
	Dd	1921	CGTGCCGAATTC 1932	
RESULT 3				
AR203450		AR203450	3969 bp DNA linear PAT 20-JUN-2002	
LOCUS		Sequence 4 from patent US 6365386.		
DEFINITION				
AUTHORS		Hoshino,T., Ojima,K. and Setoguchi,Y.		
TITLE		Astaxanthin synthase		
JOURNAL		Patent: US 6365386-A 4 02-APR-2002;		
FEATURES		Location/Qualifiers		
source		1..3969		
		/organism="unknown"		
ORIGIN		/mol_type="unassigned DNA"		
Query Match		18.3%; Score 353.2; DB 6; Length 3969;		
Best Local Similarity		99.2%; Pred. No. 3.7e-96;		
Matches 355; Conservative		0; Mismatches 3; Indels 0; Gaps 0;		
Qy	1514	TTGCTTTGGTTGGCGATTTCGTGCGCGAGATGAAGGCGTCTTGTTGTGACATCTCCG	1573	
Dd	3508	TCGCAGTGGTTGGCGATTTCGTGTCGCGAGATGAAGGCGTCTTGTTGTGACATCTCCG	3567	
Qy	1574	TCGGGTCCAGTTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACATCACCTTGATCAT	1633	
Dd	3568	TCGGGTCCAGTTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACATCACCTTGATCAT	3627	
Qy	1634	TTCCCGTCTCTCGAATCGTTGGTAGAGAAAGAGGGGTACAGATGCGTTTCGAGGTCAA	1693	
Dd	3628	TTCCCGTCTCTCGAATCGTTGGTAGAGAAAGAGGGGTACAGATGCGTTTCGAGGTCAA	3687	
Qy	1694	GCGCGTTCGAATGAGTTGATTTCTTCATATGTTTAAGAGAAAGTTCTATATCTGAGAATGTGTG	1753	
Dd	3688	GCGCGTTCGAATGAGTTGATTTCTTCATATGTTTAAGAGAAAGTTCTATATCTGAGAATGTGTG	3747	
Qy	1754	ACTAGACAATGCGCTTCTTTGTATCGAATTTGTTTCTCATACCCGGCGAGCGCTATGACT	1813	
Dd	3748	ACTAGACAATGCGCTTCTTTGTATCGAATTTGTTTCTCATACCCGGCGAGCGCTATGACT	3807	

Qy	1814	TCTAGTCGCTATCGTCGGCTCTGGACTCTCTCTTACCTATATATATATTCATCCG	1871
Db	3808	TCTAGTCGCTATCGTCGGCTCTGGACTCTCTCTTACCTATATATATTCATCCG	3865
RESULT 4			
AX034666			
LOCUS	AX034666	3969 bp	DNA
DEFINITION	Sequence 3 from Patent EP1035206.		linear
ACCESSION	AX034666		
VERSION	AX034666.1	GI:11190640	
KEYWORDS	Xanthophyllomyces dendrorhous (anamorph: Phaffia rhodozyma)		
SOURCE	Xanthophyllomyces dendrorhous		
ORGANISM	Eukaryota; Fungi; Basidiomycota; Hymenomycetes;		
	Heterobasidiomycetes; Tremellomycetidae; Cystofilobasidiales;		
	Cystofilobasidiaceae; Xanthophyllomyces.		
REFERENCE	1		
AUTHORS	Hoshino, T., Ojima, K. and Setoguchi, Y.		
TITLE	Ataxanthin synthetase		
JOURNAL	Patent: EP 1035206-A 3 13-SEP-2000;		
	KOPFMAN LA ROCHE (CH)		
FEATURES	Location/Qualifiers		
source	1..3969		
	/organism="Xanthophyllomyces dendrorhous"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:5421"		
5'UTR	517..518		
	/note="EXPERIMENTAL"		
exon	535..652		
intron	653..734		
exon	735..783		
intron	784..898		
exon	899..1015		
intron	1016..1087		
exon	1088..1179		
intron	1180..1302		
exon	1303..1517		
intron	1518..1600		
exon	1601..1634		
intron	1635..1723		
exon	1724..1866		
intron	1867..1939		
exon	1940..1999		
intron	2000..2081		
exon	2082..2181		
intron	2182..2257		
exon	2258..2354		
intron	2355..2431		
exon	2432..2542		
intron	2543..2618		
exon	2619..2652		
intron	2653..2742		
exon	2743..2814		
intron	2815..2962		
exon	2963..3050		
intron	3051..3113		
exon	3114..3171		
intron	3172..3247		
exon	3248..3321		
intron	3322..3398		
exon	3399..3423		
intron	3424..3513		
exon	3514..3700		
polyA_site	3865..3866		
	/note="EXPERIMENTAL"		
ORIGIN			
Query Match	18.3%;	Score 353.2;	DB 6; Length 3969;
Best Local Similarity	99.2%;	Pred No. 3.7e-96;	
Matches 355; Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
Qy	1514	TTGCTTTGGTGGCGAATTGCTGTCGCCGAGATGAAGGCGCTTCTTGTTGTAATCTCCG	1573



QY 1117 AGGATAAACTTCGAGAGAAATT-----TGTGATGTCGACGAGATATGCTTACGCTAG 1170  
 DB 1025 AGCAGAGATTCATGAGAGATGATTCATTTCTCCTGACGAGGCGAGTCTTACTTATG 1084  
 QY 1171 ACAGAACTTAATGCGTTCCTTATCTCGAAGCGTTTGAAGGAGTCTTCTGCTAGACC 1230  
 DB 1085 ACATTCGATCGAGATGAATATCTTATGATGTCATTCAGGAGACCTCAGGTTGATC 1144  
 QY 1231 CTCCT 1235  
 DB 1145 CCCCT 1149  
 RESULT 6  
 LOCUS AX655393 2000 bp DNA linear PAT 22-MAR-2003  
 DEFINITION Sequence 5263 from Patent WO03000898.  
 ACCESSION AX655393  
 VERSION AX655393.1 GI:29158207  
 KEYWORDS Oryza sativa  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1  
 AUTHORS Chang H.S., Chen.W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,  
 Katagiri,F., Qian,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.  
 TITLE Plant genes involved in defense against pathogens  
 JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;  
 SYNGENTA Participations AG (CH)  
 FEATURES  
 source Location/Qualifiers  
 1..2000  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 ORIGIN  
 Query Match 2.7%; Score 51.8; DB 6; Length 2000;  
 Best Local Similarity 9.8%; Pred. No. 0.0011;  
 Matches 72; Conservative 343; Mismatches 310; Indels 10; Gaps 3;  
 QY 421 ACCGATCGAAGGATCATGATCCCTCTGTCGCTCAGCGGCTTAAGTCATGCTCC 480  
 DB 48 RCGSCGCKMTTRKSKWYSASSAGRTGSKSGSYSGKMKRYKRGRGRRG 107  
 QY 481 CAATTTCTTAGAAAAAGTATGGAACCTGTGCAAGATGATGGAGATCGCGCTGAGA 540  
 DB 108 MRRSRMRWGRYRCARSGRMAGGGRMMWGGKSRMSYMMCYARGCGSKRKS--KGS 165  
 QY 541 AGGATATGCCCTGGAGAGTCGCGCGGTGAAAGAGAGGCAACAGACTCGACAGCGAG 600  
 DB 166 WGTCTRRGARGGSGWSGAKYKSGMSXRMWMMWCGRRSAYSGRYGTGRKYGTG 225  
 QY 601 GAGTCGATGAAGGATGGGTGCGTTCGAGCTACTCTGGAGCTCATGGCTTCTGCAAGAT 660  
 DB 226 KMTYISASRCMAYTTSYACSSYTWCRSKRSMMWMMKWRWSRYGWSYKWM 285  
 QY 661 TTGACTATAAGAGCGACTCGCTCGAGAACAGACCAATGAGCTATGTCGCTTTTCTCG 720  
 DB 286 MCTAYKYSYRWYRWYRGGRGATRYWGRGYMSRMAMMYKQYVRYGKMKRGWAG 345  
 QY 721 GACTTACGATGGGTTCCTCTACTCTTGGACTGCTTCAAGGCTCATATGCGGATTTG 780  
 DB 346 RMMRSMCRWSKACATYMRWMMWMTTTRRRWAKKSRSTRRRKRWKMKRYKMRYSR 405  
 QY 781 TACCTTACTTCGAACTATGAACGAGACATGAGA-----TACCTTTGACTCAAGAT 834  
 DB 406 MRSCKRMRWKRCSGRAWMGRCGTCRMKSYGMWEMKSKWEMKASKYKMSMRYWRK 465  
 QY 835 TAGCAGTTTCCGACGAGCTTGGATCGAGCTTATGAGCAAAAGACGACGCGCTGCTG 894  
 DB 466 KCSRTTWMGTRGGMMGTGRCRYKRSKMKRKRRRRGRMYMRWRYMSARYTMR 525

QY 895 GCTCAGCTTCCGATCAGGCTGTGTGATAAAAGAGATGTTCAAGGTCGGGATATCCTAAGTC 954  
 DB 526 YCARKKYSYSAARFACWYRGKYWAGWMMKRYKRYMTYKMMWYKRYKSKSWYCKWS 585  
 QY 955 TCCTAGTAGAGCAACATCGCC--GCCAACCTGCTGAATCTCAAAAGCTGTCCGATGA 1012  
 DB 586 YIASCKMSARKAGAKWCKSKMSANSKSMRSRCKRCKASKESSAKRYAMWGMGTSGSRM 645  
 QY 1013 GGAGGTACTCGCTCAGATCAGTAACCTGTATTGCTGATATGAACTTCTTCACAGT 1072  
 DB 646 SRWKSCTYWRKWSMKSTCTWYMYNSKYTYAKYSGYWRYYRAMCMYWRWYRYRYKSY 705  
 QY 1073 CTTGACATGGATGTTTCAAGGCTCTCAGAACACAAAGCCGCTCAGGATAAACTTCGAGA 1132  
 DB 706 MTYAWVTSSTRMAMTGKYSGRYWTWYKCKSKWYSMMYYWVWVWVWVWVWVWVWV 765  
 QY 1133 AGAAATTTGTCAGAT 1147  
 DB 766 TRMMWYRYSMKWY 780  
 RESULT 7  
 LOCUS HSA563377 1062 bp mRNA linear PRI 15-MAY-2003  
 DEFINITION Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).  
 ACCESSION AJ563377  
 VERSION AJ563377.1 GI:30840240  
 KEYWORDS CYP3A43/CYP3A4 gene; cytochrome P450.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Finta,C. and Zaphiropoulos,P.G.  
 TITLE Intergenic mRNA molecules resulting from trans-splicing  
 JOURNAL J. Biol. Chem. 277 (8), 5882-5890 (2002)  
 MEDLINE 21839017  
 PUBMED 11726664  
 REFERENCE 2 (bases 1 to 1062)  
 AUTHORS Zaphiropoulos,P.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska  
 Institute, Novum, 141 57 Huddinge, SWEDEN  
 FEATURES  
 source Location/Qualifiers  
 1..1062  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="liver"  
 1..1062  
 /genes="CYP3A43/CYP3A4"  
 /gene="CYP3A43/CYP3A4"  
 /codon\_start=1  
 /product="cytochrome P450"  
 /protein\_id="CAD91345.1"  
 /db\_xref="GI:30840241"  
 /translation="MDLIPNFMETWLVATSLVLLYIVFGAYSMVITSTSGVNI  
 SLNQPFFVENTKLLRDFLDLPFLSTVFPFLIPLELVNICVFFREVNFIRKS  
 VKRKSRLEDQKHVDFQLMIDQNSKETESHKALSDLELVASQIIFIAGIETT  
 SVLSFTFMYELATHPDVQOKLEIDAVLPNKAPPTDVLQWELDWVNETLRLFP  
 IANRLERVKCKDVEINGMFTPKGVVYMIYSALHRDPKYTFEPKFLPERFSKKNKN  
 IDPYVTFGSGPRNCIGMRFALNMKALIRVLQNFSPKPKETQIPLKLSIGLLQ  
 PEKPVLVKESRDGTSGA"  
 ORIGIN  
 Query Match 2.5%; Score 48.4; DB 9; Length 1062;  
 Best Local Similarity 50.7%; Pred. No. 0.01;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
 QY 997 AAAAGCTTCCGATGAGAGGATCTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056



```
Db 413 AAGCTCTGCGGATCTGGAGCTCGTGGCCCAATCAATTATCTTTATTTTGGCTGCTATG 472
QY 1057 AAACCTTTCTCGACAGCTTTGACATGGATGTTTTCACCGACTCTCAGAAAGACAAAGCCGTTTC 1116
Db 473 AAACACGAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 532
QY 1117 AGGATTAACCTTCGAGAGAAATTTTCAGATCGACACGGATATGCTTACGCTAGACGAAC 1176
Db 533 AGCAGAAATTCGAGGAGGAAATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 592
QY 1177 TTAATCGCTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCTGCTAGACC 1230
Db 593 ATACTGTGTACAGATGGAGTATCTTGACATGGTGTGATGAAGAAACGCTCAGATTATTC 652
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGATGAAGACTTCA 1278
Db 653 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 700

RESULT 8
HSA563376 1365 bp mRNA linear PRI 21-MAY-2003
LOCUS Homo sapiens partial mRNA for cytochrome P450 (CYP3A43 and CYP3A4
DEFINITION genes).
ACCESSION AJ563376
VERSION AJ563376.1 GI:30840238
KEYWORDS CYP3A4 gene; CYP3A43 gene; cytochrome P450.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Finta,C. and Zaphiropoulos,P.G.
TITLE Intergenic mRNA molecules resulting from trans-splicing
JOURNAL J. Biol. Chem. 277 (8), 5882-5890 (2002)
MEDLINE 2183017
PubMed 11728664
REFERENCE 2 (bases 1 to 1365)
AUTHORS Zaphiropoulos,P.G.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska
Institute, Novum, 141 57 Huddinge, SWEDEN

FEATURES
source
1. .1365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
1 71
/genes="CYP3A43"
1. .>71
/genes="CYP3A43"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAD91644.1"
/db_xref="GI:31043862"
/translation="MDLIPFAMETWLVATSLVLLY"
72..1365
/genes="CYP3A4"
<72..1365
/genes="CYP3A4"
/codon_start=2
/product="cytochrome P450"
/protein_id="CAD91645.1"
/db_xref="GI:31043863"

gene
1. .1365
/genes="CYP3A4"
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
1 71
/genes="CYP3A43"
1. .>71
/genes="CYP3A43"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAD91644.1"
/db_xref="GI:31043862"
/translation="MDLIPFAMETWLVATSLVLLY"
72..1365
/genes="CYP3A4"
<72..1365
/genes="CYP3A4"
/codon_start=2
/product="cytochrome P450"
/protein_id="CAD91645.1"
/db_xref="GI:31043863"

gene
1. .1365
/genes="CYP3A4"
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
1. .1512
/product="cytochrome P4503A4"
```

```
ORIGIN
Query Match 2.5%; Score 48.4; DB 9; Length 1365;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTCTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056
Db 716 AAGCTCTCTCGATCTGGAGCTCGTGGCCCAATCAATTATCTTTATTTGCTGGCTATG 775
QY 1057 AAACCTTTCTCGACAGTCTTTGACATGGATGTTTTCACCGACTCTCAGAAAGACAAAGCCGTTTC 1116
Db 776 AAACACGAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835
QY 1117 AGGATTAACCTTCGAGAGAAATTTTCAGATCGACACGGATATGCTTACGCTAGACGAAC 1176
Db 836 AGCAGAAATTCGAGGAGGAAATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 895
QY 1177 TTAATCGCTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCTGCTAGACC 1230
Db 896 ATACTGTGTACAGATGGAGTATCTTGACATGGTGTGATGAAGAAACGCTCAGATTATTC 955
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGATGAAGACTTCA 1278
Db 956 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1003

RESULT 9
E10636 1512 bp RNA linear PAT 29-SEP-1997
LOCUS Human cDNA encoding cytochrome P4503A4.
DEFINITION E10636
ACCESSION E10636.1 GI:22027692
VERSION JP 1996027197-A/5.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1512)
AUTHORS Funae,Y., Imaoka,S., Matsuki,Y., Hayashi,K. and Yabusaki,Y.
TITLE ANTIBODY RECOGNIZING CYTOCHROME P4502A4 ORIGINATED FROM MAN
JOURNAL Patent: JP 1996027197-A 5 30-JAN-1996;
SUMITOMO CHEM. CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996027197-A/5
PD 30-JAN-1996
PF 13-JUL-1994 JP 1994161552
PI FUNAE YOSHIHIKO, IMAOKA SUSUMU, MATSUKI YASUSHI, HAYASHI KOJI,
PI YABUSAKI YOSHIYASU
PC C07K16/18,C12N15/09//C12N1/19,C12N9/02,G01N33/53,(C12N1/19, PC
C12R1:865);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1512
FT /organism="Homo sapiens"
FT /tissue_type="liver"
FT 1..1512
FT CDS /product="cytochrome P4503A4"
FT
FEATURES
source
1..1512
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTCTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056
```

```

Db      863 AAGCTCTGTCGAGCTCGTGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057 AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAAAGACAAAGCCGTTTC 1116
Db      923 AAACACGAGCAGTGTCTCTCTCTCAATATGATGAACCTGGCCACTCACCCTGATGTCC 982
QY      1117 AGGATAAACTTCGAGAGAAATTTGTACAGTCGACACGGATATGCCCTACGCTAGACGAAC 1176
Db      983 ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCTATG 1042
QY      1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
Db      1043 ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db      1043 ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db      1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1150

RESULT 10
LOCUS      E10855      1512 bp      RNA      linear      PAT 29-SEP-1997
DEFINITION CDNA encoding human cytochrome P450.
ACCESSION  E10855
VERSION     E10855.1 GI:22027949
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
          Hayashi,K., Sakaki,T., Yabueaki,Y., Komai,K., Kaneko,H. and
          Nakatsuka,I.
TITLE      METHOD FOR EVALUATING SAFETY
JOURNAL    Patent: JP 1996056695-A 4 05-MAR-1996;
COMMENT    SUMITOMO CHEM CO LTD
          OS Homo sapiens (human)
          PN JP 1996056695-A/4
          PD 05-MAR-1996
          PF 15-JUL-1994 JP 1994164184
          PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
          17-JUN-1994 JP 94P 136053
          PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
          KOICHIRO,
          P1 KANEKO HIDEO, NAKATSUKA IWAO
          PC C12Q1/02,C12M1/34,C12Q1/26;
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          FH Key Location/Qualifiers
          FT source 1..1512
          FT CDS 1..1512
          FT /organism='Homo sapiens'
          FT /product='human cytochrome p450 3A4'.
          FT Location/Qualifiers
          FT source 1..1512
          FT /organism='Homo sapiens'
          FT /mol_type='genomic RNA'
          FT /db_xref='taxon:9606'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGAGTGGAGGTTACTCGTTCAGATCAGTAACTGTTATTTCTGGATG 1056
Db 863 AAGCTCTGTCGAGCTCGTGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY 1057 AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAAAGACAAAGCCGTTTC 1116
Db 923 AAACACGAGCAGTGTCTCTCTCTCAATATGATGAACCTGGCCACTCACCCTGATGTCC 982
QY 1117 AGGATAAACTTCGAGAGAAATTTGTACAGTCGACACGGATATGCCCTACGCTAGACGAAC 1176
Db 983 ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCTATG 1042
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
Db 1043 ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY 1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1150

RESULT 12
LOCUS      HSA563375      1512 bp      mRNA      linear      PRI 15-MAY-2003
DEFINITION Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION  AJ563375
VERSION     AJ563375.1 GI:30840236
KEYWORDS   CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

```

```

Db      923 AAACACGAGCAGTGTCTCTCTCTCAATATGATGAACCTGGCCACTCACCCTGATGTCC 982
QY      1117 AGGATAAACTTCGAGAGAAATTTGTACAGTCGACACGGATATGCCCTACGCTAGACGAAC 1176
Db      983 ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCTATG 1042
QY      1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
Db      1043 ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db      1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1150

RESULT 11
LOCUS      AR399323      1512 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 4 from patent US 6620593.
ACCESSION  AR399323
VERSION     AR399323.1 GI:40141119
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1512)
AUTHORS   Hayashi,K., Sakaki,T., Yabueaki,Y., Komai,K., Kaneko,H. and
          Nakatsuka,I.
TITLE      Method for safety evaluation of chemical compound using recombinant
          yeast expressing human cytochrome P450
JOURNAL    Patent: US 6620593-A 4 16-SEP-2003;
FEATURES   Location/Qualifiers
            1..1512
            /organism='unknown'
            /mol_type='genomic DNA'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGAGTGGAGGTTACTCGTTCAGATCAGTAACTGTTATTTCTGGATG 1056
Db 863 AAGCTCTGTCGAGCTCGTGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY 1057 AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAAAGACAAAGCCGTTTC 1116
Db 923 AAACACGAGCAGTGTCTCTCTCTCAATATGATGAACCTGGCCACTCACCCTGATGTCC 982
QY 1117 AGGATAAACTTCGAGAGAAATTTGTACAGTCGACACGGATATGCCCTACGCTAGACGAAC 1176
Db 983 ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCTATG 1042
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
Db 1043 ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY 1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1150

RESULT 12
LOCUS      HSA563375      1512 bp      mRNA      linear      PRI 15-MAY-2003
DEFINITION Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION  AJ563375
VERSION     AJ563375.1 GI:30840236
KEYWORDS   CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

```

## REFERENCE

1 Finta, C. and Zaphiropoulos, P.G.  
TITLE Interfering mRNA molecules resulting from trans-splicing  
JOURNAL J. Biol. Chem. 277 (8), 5882-5890 (2002)  
MEDLINE 21839017  
PUBMED 11726564

## REFERENCE

2 Zaphiropoulos, P.G.  
TITLE Direct Submission  
JOURNAL Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska  
Institute, Novum, 141 57 Huddinge, SWEDEN

## FEATURES

Location/Qualifiers

source

1. .1512

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="liver"

1. .1512

/gene="CYP3A43/CYP3A4"

1. .1512

/gene="CYP3A43/CYP3A4"

/codon\_start=1

/product="cytochrome P450"

/protein\_id="CAD91343.1"

/db\_xref="GI:30840237"

/translation="MDLIPNFMETWLVATSLVLLIYXTHSHGLFKLIGPPTPL

/translaton="MDLIPNFMETWLVATSLVLLIYXTHSHGLFKLIGPPTPL

PFLNLSYHKGFCWMECHKYKVMGQVQVLAITDPMIKTLVLEKCVSF

TNRPPGPGVPMKSAISIADEEWKRLRSLLSTFTSGKLMVPIIAQVGVLRNL

PREATGKPTLKDVFAYSMVDUTSFGVNIDSLNPDQPPVNTKLLRFDLDP

FFLSITVFFPILEVLNLCVFPRENTVLRKSVKMKESLREDQKRVDFLOMI

IDVQNKETESHKALSDLELVAQSLIFIFAGYETTSVLSFIMELAHSDVQKCEZ

ISDALNKAPTDTVLQMEYLDVNVNETLRLFIAMRLVCKKQVDEINGMFIKGV

VVMIPSVLHHRDPKWTPEKFLPERFSKKNIDPIYIYTFPGSPRCIGNRFALM

NMKLALIRVLQNFSEKPKETQIPLKLSGLLQPKPVVLKVESRDGTVSGA"

## ORIGIN

Query Match 2.5%; Score 48.4; DB 9; Length 1512;

Best Local Similarity 50.7%; Pred. No. 0.011;

Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056

Db 863 AAGCTCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 922

QY 1057 AAACCTTCTTCGACAGTCTTGATGGATGTTTCCAGCATCTCAGAGCAAAAGCGGTTTC 1116

Db 923 AAACCAAGCAGCAGTCTTCTCTCTTCAATATGATGAACTGGCCACTCACCCCTGATGTC 982

QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTTACGCTAGACGAAC 1176

Db 983 AGCAGAAATCTGAGGAGAAATTTGATGAGTTTACCCATTAAGGACACCCACCTATG 1042

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTTCTGCTAGACC 1230

Db 1043 ATACTGTGCTACAGATGAGTATCTTTGATGATGGTGGTGAATGAACGCTCAGATTATTC 1102

QY 1231 CTCCTAGTCGATGCTTACCGTGAATCTTAAAGGATGAGACTTCA 1278

Db 1103 CAATTGCTATGAGATTGAGAGGGTCTGCAAAAAGATGTTGATATCA 1150

## RESULT 13

AX408939

LOCUS

DEFINITION

AX408939

AX408939

AX408939.1

GI:21441644

KEYWORDS

ORGANISM

source

1

AX408939

Sequence 1586 from Patent WO029103.

AX408939

AX408939.1

GI:21441644

KEYWORDS

ORGANISM

source

1

## AUTHORS

Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.  
TITLE Gene expression profiles in liver cancer  
JOURNAL Patent: WO 029103-A 1586 11-APR-2002;  
GENE LOGIC INC (US)

## FEATURES

Location/Qualifiers

source

1. .2011

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

/note="EMBL/GenBank Accession No. D00003"

## ORIGIN

Query Match 2.5%; Score 48.4; DB 6; Length 2011;

Best Local Similarity 50.7%; Pred. No. 0.012;

Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056

Db 926 AAGCTCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 985

QY 1057 AAACCTTCTTCGACAGTCTTGATGGATGTTTCCAGCATCTCAGAGCAAAAGCGGTTTC 1116

Db 986 AAACCAAGCAGCAGTCTTCTCTCTTCAATATGATGAACTGGCCACTCACCCCTGATGTC 1045

QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTTACGCTAGACGAAC 1176

Db 1046 AGCAGAAATCTGAGGAGAAATTTGATGAGTTTACCCATTAAGGACACCCACCTATG 1105

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTTCTGCTAGACC 1230

Db 1106 ATACTGTGCTACAGATGAGTATCTTTGATGATGGTGGTGAATGAACGCTCAGATTATTC 1165

QY 1231 CTCCTAGTCGATGCTTACCGTGAATCTTAAAGGATGAGACTTCA 1278

Db 1166 CAATTGCTATGAGATTGAGAGGGTCTGCAAAAAGATGTTGATATCA 1213

## RESULT 14

HUMCYPHLP

LOCUS

DEFINITION

Human liver glucocorticoid-inducible cytochrome P-450 (HLP) mRNA,

complete cds.

ACCESSION

M13785

VERSION

1

KEYWORDS

cytochrome; cytochrome P450; microsomal hemoprotein.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2011)

AUTHORS

Molowa, D.T., Schuetz, E.G., Wrighton, S.A., Watkins, P.B., Kremers, P.,

Mender-Picon, G., Parker, G.A. and Guzelian, P.S.

Complete cDNA sequence of a cytochrome P-450 inducible by

glucocorticoids in human liver

JOURNAL

MEDLINE

PUBMED

COMMENT

Original source text: Human (adult) liver, cDNA to mRNA, clones

HP[55.12-9].

Draft entry and sequence in computer-readable form for [1] kindly

provided by D.T.Molowa, 19-NOV-1986.

The human genome probably contains multiple genes or pseudogenes

related to glucocorticoid-inducible cytochrome P-450.

## FEATURES

Location/Qualifiers

source

1. .2011

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

61. .1575

/note="glucocorticoid-inducible cytochrome P-450"

/codon\_start=1

/protein\_id="AAA35742.1"

/db\_xref="GI:181354"

## CDS

/translation="MALIPDLAMETWLLAVSLVLYLGYTHSHGLFKKLGIPGPTEL  
PFLGNILSYHKGFCMFMECHKYKGVYDQGPVLAITDPDMIKLVLYKECVSVF  
INREPTGPGVGMKSAISIAEDENWRKLSLPTFTSGKKNWVPIAQIGDVLVRLN  
RRERETGPVTKDVFAGYMDVITSSFGVNDLNNPQDPLVENTKLLRFDLDP  
PFLSITVPPFLIPLEVLNLCVFPREVTNFKAVKMKESRLEDTQKRVDFLOLMI  
DSHKNSKTESHKSLSLVAQSIIFIFAGYETTSVLSTFMYSLACHPDVQOQLQE  
EIDAVLPNKAPTDYTLQMEYLDWVNETLRFPIAMRLERVCKDKVEINGMFIKPG  
WVMVPSVALHRDPKYWTEPEKFLPERFSKKNNDIPYITPFGSGPRNCIGMRPAL  
MMKMLALIRVLQNSFPKCKETQIPKLKSLGLLQPEKFPVLKVESRDGTVSGA"

ORIGIN 170 bp upstream of EcoRI site.

Query Match 2.5%; Score 48.4; DB 9; Length 2011;  
Best Local Similarity 50.7%; Pred. No. 0.012;  
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGACGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG 1056  
DB 926 AAGCTGTGCCACTGAGAGCTGCGGCCCAATCAATATTTATTTTGTGGCTATG 985  
QY 1057 AAAGCTGTCGACGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG 1116  
DB 986 AAAGCTGTCGACGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG 1045  
QY 1117 AAGATAAAGCTGTCGACGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG 1176  
DB 1046 AAGATAAAGCTGTCGACGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG 1105  
QY 1177 TTAATGCGGTGC-----CTTATCTCGAAGCGTGTGTTAAGGAGTCTCTCGTCTAGACC 1230  
DB 1106 ATACTGTCTCAGATGAGGATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1165  
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278  
DB 1166 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA 1213

RESULT 15  
LOCUS HUMF450M 2011 bp mRNA linear PRI 13-NOV-1997  
DEFINITION Homo sapiens mRNA for cytochrome P-450, complete cds.  
ACCESSION D00003 N00003  
VERSION D00003.1 GI:219962  
KEYWORDS cytochrome P-450.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2011)  
Molowa,D.T., Schuetz,E.G., Wrighton,S.A., Watkins,P.B., Krewers,P.,  
Mendez-Picon,G., Parker,G.A. and Guzelian,P.S.  
Complete cDNA sequence of a cytochrome P-450 inducible by  
glucocorticoids in human liver  
Proc. Natl. Acad. Sci. U.S.A. 83 (14), 5311-5315 (1986)  
JOURNAL MEDLINE 86259780  
PUBMED 3460094  
COMMENT The liver cytochromes P-450 are supergene family of microsomal  
hemoproteins that catalyze the oxidative biotransformation of  
numerous endogenous substrates as well as xenobiotics. They used a  
monoclonal antibody that recognizes purified human liver cytochrome  
P-450 (Hlp) to screen a human liver cDNA library in lambda-gt10.  
They isolated and sequenced two overlapping cDNA clones that span  
the entire 2011 bases of an mRNA that codes for a protein of 504  
amino acids. Hlp mRNA shares more than 70% sequence homology with  
related proteins from the rat and rabbit but less than 40% homology  
with other published cytochrome P-450 genes. With a  
glucocorticoid there was a several fold induction of two mRNA  
species hybridizable to Hlp cDNAs as well as an induction of  
immunoreactive Hlp protein. They conclude that Hlp is a human  
representative of a conserved glucocorticoid-inducible cytochrome  
P-450 gene family whose mechanism of induction involves  
accumulation of Hlp mRNA.

FEATURES  
source Location/Qualifiers  
1..2011

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/issue\_type="liver"  
/clone\_lib="human liver cDNA library provided from Meloy  
Laboratories"  
/notes="169 bp upstream of EcoRI site"  
61..1575  
/codon\_start=1  
/product="cytochrome P-450"  
/protein\_id="BAA00001.1"  
/db\_xref="GI:219962"  
/translation="MALIPDLAMETWLLAVSLVLYLGYTHSHGLFKKLGIPGPTEL  
PFLGNILSYHKGFCMFMECHKYKGVYDQGPVLAITDPDMIKLVLYKECVSVF  
INREPTGPGVGMKSAISIAEDENWRKLSLPTFTSGKKNWVPIAQIGDVLVRLN  
RRERETGPVTKDVFAGYMDVITSSFGVNDLNNPQDPLVENTKLLRFDLDP  
PFLSITVPPFLIPLEVLNLCVFPREVTNFKAVKMKESRLEDTQKRVDFLOLMI  
DSHKNSKTESHKSLSLVAQSIIFIFAGYETTSVLSTFMYSLACHPDVQOQLQE  
EIDAVLPNKAPTDYTLQMEYLDWVNETLRFPIAMRLERVCKDKVEINGMFIKPG  
WVMVPSVALHRDPKYWTEPEKFLPERFSKKNNDIPYITPFGSGPRNCIGMRPAL  
MMKMLALIRVLQNSFPKCKETQIPKLKSLGLLQPEKFPVLKVESRDGTVSGA"

ORIGIN

Query Match 2.5%; Score 48.4; DB 9; Length 2011;  
Best Local Similarity 50.7%; Pred. No. 0.012;  
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
QY 997 AAAGCTGTCGACGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG 1056  
DB 926 AAGCTGTGCCACTGAGAGCTGCGGCCCAATCAATATTTATTTTGTGGCTATG 985  
QY 1057 AAAGCTGTCGACGAGTACTCGCTCAGATGAGTAACTGTTATTTGCTGGATATG 1116  
DB 986 AAAGCTGTCGACGAGGAGTACTCGCTCAGATGAGTAACTGTTATTTGCTGGATATG 1045  
QY 1117 AAGATAAAGCTGTCGACGAGGAGTACTCGCTCAGATGAGTAACTGTTATTTGCTGGATATG 1176  
DB 1046 AAGATAAAGCTGTCGACGAGGAGTACTCGCTCAGATGAGTAACTGTTATTTGCTGGATATG 1105  
QY 1177 TTAATGCGGTGC-----CTTATCTCGAAGCGTGTGTTAAGGAGTCTCTCGTCTAGACC 1230  
DB 1106 ATACTGTCTCAGATGAGGATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1165  
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278  
DB 1166 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA 1213

Search completed: April 6, 2004, 16:19:22  
Job time : 7671 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

\*OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 08:05:34 ; Search time 833 Seconds

(without alignments)  
9852.964 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932  
Sequence: 1 gaattcgacagagccacc.....agcggctcgtcggaattc 1932

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

-Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002s: \*  
7: Geneseqn2003as: \*  
8: Geneseqn2003bs: \*  
9: Geneseqn2003cs: \*  
10: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1932	100.0	1932	3	AA64472
2	353.2	18.3	3969	3	AA64473
3	51.8	2.7	2000	7	ADA71938
4	49.4	2.6	1501	6	ABL89830
5	48.4	2.5	991	6	ABK97688
6	48.4	2.5	1062	6	ABK97694
7	48.4	2.5	1365	6	ABK97693
8	48.4	2.5	1512	2	AAQ87717
9	48.4	2.5	1512	2	AAT17399
10	48.4	2.5	1512	2	AAT28383
11	48.4	2.5	1512	6	ABK97692
12	48.4	2.5	1512	6	ACF06054
13	48.4	2.5	2011	6	ABN95088
14	48.4	2.5	2759	2	AAK28235
15	48.4	2.5	2764	7	ABX77172
16	48.4	2.5	2849	6	ABN95612
17	48.2	2.5	312	7	ABX39446
18	48.2	2.5	1645	3	AAC77026
19	47.6	2.5	3059	3	AAC99040
20	47.4	2.5	2160	2	AAZ10641
21	47	2.4	445	5	ABV39053
22	46.8	2.4	1712	2	AAV34249
23	46.8	2.4	1712	7	ACD08120

24	46.8	2.4	1822	2	AAV34248	AAV34248 Human sec
25	46.8	2.4	1822	7	ACD08119	ACD08119 cDNA enco
26	46.8	2.4	6004	9	ADB58866	ADB58866 Toxicity-
27	46.4	2.4	1552	4	ABL14451	ABL14451 Drosophil
28	45.4	2.3	576	5	ABV44839	ABV44839 Human pro
29	45.2	2.3	413	7	ABK36570	ABK36570 Bovine ES
30	45.2	2.3	1608	6	ABK97643	ABK97643 DNA enco
31	45.2	2.3	1608	9	ADD29816	ADD29816 Human tum
32	45.2	2.3	1971	6	ABL62314	ABL62314 Colon ade
33	45.2	2.3	1971	6	ABL64874	ABL64874 Lung canc
34	45.2	2.3	1971	6	ABN95091	ABN95091 Gene #158
35	45.2	2.3	1971	6	ABN95091	ABN95091 Gene #158
36	45.2	2.3	2010	3	AAK36214	AAK36214 Human cyt
37	45.2	2.3	2010	3	AAK79851	AAK79851 Human sec
38	45.2	2.3	2379	4	AAH57474	AAH57474 Human liv
39	45.2	2.3	4467	4	AAH72650	AAH72650 Human cer
40	45	2.3	428	5	ABV56764	ABV56764 Human pro
41	45	2.3	1831	2	AAK60780	AAK60780 Soybean c
42	44.8	2.3	1801	3	AAZ35893	AAZ35893 cDNA enco
43	44.8	2.3	2878	3	AAZ93078	AAZ93078 Partial s
44	44.8	2.3	4467	5	ABV22264	ABV22264 Human pro
45	44.8	2.3	4467	5	ABV28101	ABV28101 Human pro
46	44.6	2.3	441	5	ABV44692	ABV44692 Human pro

ALIGNMENTS

RESULT 1  
AAA64472  
ID AAA64472 standard; cDNA; 1932 BP.  
XX  
AC AAA64472;  
XX  
DT 15-SEP-2003 (revised)  
DT 02-JAN-2001 (first entry)  
XX  
DE cDNA encoding an astaxanthin synthetase polypeptide.  
XX  
KW Astaxanthin synthetase; astaxanthin; beta-carotene; carotenogenic yeast;  
KW antioxidant; cancer; colouring reagent; farmed fish; salmon; ss.  
XX  
OS Xanthophyllomyces dendrorhous.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 33..1706  
FT /\*tag= a  
FT /product= "astaxanthin synthetase"  
FT polyA\_signal  
FT 1871  
FT /\*tag= b  
XX  
EP1035206-Al.  
XX  
PD 13-SEP-2000.  
XX  
PF 03-MAR-2000; 2000EP-00104430.  
XX  
PR 09-MAR-1999; 99EP-00104668.  
PR 01-FEB-2000; 2000EP-00101656.  
XX  
PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
XX  
PI Hoshino T, Ojima K, Setoguchi Y;  
XX  
DR WPI; 2000-559874/52.  
XX  
P-PSDB; AAB08713.  
XX  
XX Novel polynucleotide encoding astaxanthin synthase useful for producing  
XX recombinant cells for producing astaxanthin from beta-carotene.  
XX Claim 3; Page 24-27; 46pp; English.  
XX The present sequence encodes an astaxanthin synthetase polypeptide of  
XX Pfaffia rhodozyma. The enzyme is involved in the last step of the

CC	astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.
CC	rhodzyma is a carotenogenic yeast strain. The astaxanthin synthetase
CC	polynucleotides and polypeptides are useful for producing astaxanthin.
CC	Astaxanthin is an antioxidant which may be used to protect living cells
CC	against diseases such as cancer. Astaxanthin is also used as a colouring
CC	reagent, e.g. in farmed fish like salmon to impart an orange-red
CC	coloration. (Updated on 15-SEP-2003 to standardise OS field)
XX	
SQ	Sequence 1932 BP; 495 A; 456 C; 481 G; 500 T; 0 U; 0 Other;
	Query Match 100.0%; Score 1932; DB 3; Length 1932;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1 GAAATCGGCACAGAGCGCACTACTTTCTCCATATGTTTCATCTTGGTCTTGCTCACAGGTG 60
Db	1 GAAATTCGGACAGAGCCACCTACTTTCTCCATATGTTTCATCTTGGTCTTGCTCACAGGTG 60
Qy	61 CTTTAGCGCTGGCTGCTTTCTTCATGGGCATCCATAGCGTTCTTTCAGTCTTTTACCTCGCTC 120
Db	61 CTTTAGCGCTGGCTGCTTTCTTCATGGGCATCCATAGCGTTCTTTCAGTCTTTTACCTCGCTC 120
Qy	121 CGAGGCGATCTTCACTGTATTAACCTTTAGGCGCCCGAATCATACCACTACTTTTACAGGCA 180
Db	121 CGAGGCGATCTTCACTGTATTAACCTTCAGGCGCCCGAATCATACCACTACTTTTACAGGCA 180
Qy	181 ATTTTTTTAGACATCTCTCTCAGCTCGTACAGGTGAAGAGCATGCGAAGTACAGAGAAAT 240
Db	181 ATTTTTTTAGACATCTCTCTCAGCTCGTACAGGTGAAGAGCATGCGAAGTACAGAGAAAT 240
Qy	241 ACGGAAGCAACCTTCGCGTTTGCTGGGATTCGCTGGAGCACCCGCTTTGAACTCGACCGATC 300
Db	241 ACGGAAGCAACCTTCGCGTTTGCTGGGATTCGCTGGAGCACCCGCTTTGAACTCGACCGATC 300
Qy	301 CGAAAGTCTTCAACCATGTGATGAAGAGAGCCTACGACTATCCGAAACCTGTTATGGCGG 360
Db	301 CGAAAGTCTTCAACCATGTGATGAAGAGAGCCTACGACTATCCGAAACCTGTTATGGCGG 360
Qy	361 CTCGAGTGTCTCAGAAATTGCTACCGGAGATGCTGTTGTTTACGGCGGAAGGTGAAGCTCAT 420
Db	361 CTCGAGTGTCTCAGAAATTGCTACCGGAGATGCTGTTGTTTACGGCGGAAGGTGAAGCTCAT 420
Qy	421 AGCGACATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGGTCC 480
Db	421 AGCGACATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGGTCC 480
Qy	481 CAATTTTCTTAGAAAAGGTATGGAACCTTGTTCGAAGAATGATGGAGGATGCGGCTGAGA 540
Db	481 CAATTTTCTTAGAAAAGGTATGGAACCTTGTTCGAAGAATGATGGAGGATGCGGCTGAGA 540
Qy	541 AGGATATGCCCCTGGGAGAGTGGCCGCTGAAAAGAGGCAACAGACTCGAGACCGAAG 600
Db	541 AGGATATGCCCCTGGGAGAGTGGCCGCTGAAAAGAGGCAACAGACTCGAGACCGAAG 600
Qy	601 GAGTCGATCTAAAGNATTGGGTTCGGTCAGACTACTCTTGGCGTCATGGCTCTTTCAGAGT 660
Db	601 GAGTCGATCTAAAGNATTGGGTTCGGTCAGACTACTCTTGGCGTCATGGCTCTTTCAGAGT 660
Qy	661 TTGACTATAAGAGCGACTCGCTCCAGAACAGAGCAATGAGCTCTATGTCCGCTTTTGTGCG 720
Db	661 TTGACTATAAGAGCGACTCGCTCCAGAACAGAGCAATGAGCTCTATGTCCGCTTTTGTGCG 720
Qy	721 GACTTACCGATGGGTTTGCTCTACTTTGGACTCGTTTCAGGCTATCATGTGGGATTTTG 780
Db	721 GACTTACCGATGGGTTTGCTCTACTTTGGACTCGTTTCAGGCTATCATGTGGGATTTTG 780
Qy	781 TACTCTTACTTCCGAACATAATGAACCGAGACATGAGATACCTTTGACTCAAGAGTTACGAG 840
Db	781 TACTCTTACTTCCGAACATAATGAACCGAGACATGAGATACCTTTGACTCAAGAGTTACGAG 840
Qy	841 TTTTCCCGACAGTTGGGATTCGAGCTTTATGAGAGCAAAAGAGACGAGCCGCTCTTGGCTCAG 900
Db	841 TTTTCCCGACAGTTGGGATTCGAGCTTTATGAGAGCAAAAGAGACGAGCCGCTCTTGGCTCAG 900

Qy	901	CTTCCGATCAGGCTGTTGATATAAAAGCATGTCTCAAGTCGGGATATCCTTAAGTCTCCTAG	960
Db	901	CTTCCGATCAGGCTGTTGATATAAAAGCATGTCTCAAGTCGGGATATCCTTAAGTCTCCTAG	960
Qy	961	TGAGAGCAAAATCAGCGCGCAACCTCGCTGAAATCTCAAAAGCTGTCGGATGAGGAGGTAC	1020
Db	961	TGAGAGCAAAATCAGCGCGCAACCTCGCTGAAATCTCAAAAGCTGTCGGATGAGGAGGTAC	1020
Qy	1021	TCGCTCAGATCAGTAACTGTTATTTGCTGGATATGAAACTTCTTCGACAGTCTTCACAT	1080
Db	1021	TCGCTCAGATCAGTAACTGTTATTTGCTGGATATGAAACTTCTTCGACAGTCTTCACAT	1080
Qy	1081	GGATGTTTCAACCACCTCTCAGAAAGCAAAAGCGGTTCAGGATAAACTTCGAGAGAAATTT	1140
Db	1081	GGATGTTTCAACCACCTCTCAGAAAGCAAAAGCGGTTCAGGATAAACTTCGAGAGAAATTT	1140
Qy	1141	GTCAAGTCGACGAGATATGCTTACGCTAGACGAACCTTAATCGTTGCGCTTATCTCGAAG	1200
Db	1141	GTCAAGTCGACGAGATATGCTTACGCTAGACGAACCTTAATCGTTGCGCTTATCTCGAAG	1200
Qy	1201	CGTTTGTTAAGGAGTCTCTTCGTTCTAGACCCCTCTAGTCCGTATGCTAAACCGTGAATGCT	1260
Db	1201	CGTTTGTTAAGGAGTCTCTTCGTTCTAGACCCCTCTAGTCCGTATGCTAAACCGTGAATGCT	1260
Qy	1261	TAAAGGATGAGACTTCATCCCACTTCGCGAGCGCTGTCATTGGTCAGATGGTCCGTCAC	1320
Db	1261	TAAAGGATGAGACTTCATCCCACTTCGCGAGCGCTGTCATTGGTCAGATGGTCCGTCAC	1320
Qy	1321	TCAACGAGTCCGGATCAGAAAGAACGATGGTCTATGCTTCGTTGTTCAACATCAATC	1380
Db	1321	TCAACGAGTCCGGATCAGAAAGAACGATGGTCTATGCTTCGTTGTTCAACATCAATC	1380
Qy	1381	GTTCAAGTTCATTTATGGAGAGTGCAGAGAAATTCAGACCGGAGAGGTGCTTGAGG	1440
Db	1381	GTTCAAGTTCATTTATGGAGAGTGCAGAGAAATTCAGACCGGAGAGGTGCTTGAGG	1440
Qy	1441	ACGTAAACAGACTCGCTCAACAGTATTGAAGCAACCTATGGACACCGGAGGCTTATCT	1500
Db	1441	ACGTAAACAGACTCGCTCAACAGTATTGAAGCAACCTATGGACACCGGAGGCTTATCT	1500
Qy	1501	CTGACCCAGAGCTTCGTTGGTTGGGATTTGCTGTCGCGAGATGAAGGCTTCCTGT	1560
Db	1501	CTGACCCAGAGCTTCGTTGGTTGGGATTTGCTGTCGCGAGATGAAGGCTTCCTGT	1560
Qy	1561	TTGTCACTCCGTCGGGTCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGACACA	1620
Db	1561	TTGTCACTCCGTCGGGTCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGACACA	1620
Qy	1621	TCACCTTGATCATTTCCCGTCTCGAATCGTTGGTAGAGAGAGGAGGTACCATGC	1680
Db	1621	TCACCTTGATCATTTCCCGTCTCGAATCGTTGGTAGAGAGAGGAGGTACCATGC	1680
Qy	1681	GTTTGCAGTCAAGCCGCTGCAATGAGTTGATCTTCATATGTTAAAGAAAGTTCATAT	1740
Db	1681	GTTTGCAGTCAAGCCGCTGCAATGAGTTGATCTTCATATGTTAAAGAAAGTTCATAT	1740
Qy	1741	CTGAGATGTGCATAGGACATGCTCTTTGTTATCGATTTGTTTCTCATACCCGGGC	1800
Db	1741	CTGAGATGTGTGACATAGGACATGCTCTTTGTTATCGATTTGTTTCTCATACCCGGGC	1800
Qy	1801	AGGCGCTATGACTTCTACGTCGCTATCGCTCGCTCGGACTCTCTTCTTACCTATAT	1860
Db	1801	AGGCGCTATGACTTCTACGTCGCTATCGCTCGCTCGGACTCTCTTCTTACCTATAT	1860
Qy	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCGCGCT	1920
Db	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCGCGCT	1920
Qy	1921	CGTGCGAAATTC	1932
Db	1921	CGTGCGAAATTC	1932



PT Novel polynucleotide encoding astaxanthin synthase useful for producing  
PT recombinant cells for producing astaxanthin from beta-carotene.

XX Claim 4; Page 30-33; 46pp; English.

XX The present sequence encodes an astaxanthin synthetase polypeptide of  
CC Phaffia rhodomyces. The enzyme is involved in the last step of the  
CC astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.  
CC rhodomyces is a carotenogenic yeast strain. The astaxanthin synthetase  
CC polynucleotides and polypeptides are useful for producing astaxanthin.  
CC Astaxanthin is an antioxidant which may be used to protect living cells  
CC against diseases such as cancer. Astaxanthin is also used as a colouring  
CC reagent, e.g. in farmed fish like salmon to impart an orange-red  
CC coloration. (Updated on 15-SEP-2003 to standardise OS field)

XX SQ Sequence 3969 BP; 937 A; 942 C; 869 G; 1221 T; 0 U; 0 Other;

Query Match 18.3%; Score 353.2; DB 3; Length 3969;  
Best Local Similarity 99.2%; Pred. No. 3.8e-87;  
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1514 TTGCTTTGGTTGGGATTTGCTGTCGCGAGATGAGGCTTCTTTGTCACCTCTCCG 1573  
DB 3508 TCGCAGTGTGTGGGATTTGCTGTCGCGAGATGAGGCTTCTTTGTCACCTCTCCG 3567  
QY 1574 TGGGTCCAGTTTCGAGCCCATCATCTCTCATCCAGATGACGACATCACCITGATCAT 1633  
DB 3568 TGGGTCCAGTTTCGAGCCCATCATCTCTCATCCAGATGACGACATCACCITGATCAT 3627  
QY 1634 TTCCCGTCTCGAATCGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1693  
DB 3628 TTCCCGTCTCGAATCGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3687  
QY 1694 GCGGTCGAATGAGTTGATCTTCTATATGTTAAGAGAGTTCTATATCTGAGATGTGTG 1753  
DB 3688 GCGGTCGAATGAGTTGATCTTCTATATGTTAAGAGAGTTCTATATCTGAGATGTGTG 3747  
QY 1754 ACTAGGACAAATGCTTTCTTTGATCGATTTGTTCTCATCCCGGCGAGCGGTATGACT 1813  
DB 3748 ACTAGGACAAATGCTTTCTTTGATCGATTTGTTCTCATCCCGGCGAGCGGTATGACT 3807  
QY 1814 TCTACGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1871  
DB 3808 TCTACGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3865

RESULT 3

ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

PD 03-JAN-2003.

XX 22-JUN-2001; 2001WO-1B001105.

XX 22-JUN-2001; 2001WO-1B001105.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

XX Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 2.7%; Score 51.8; DB 7; Length 2000;  
Best Local Similarity 9.8%; Pred. No. 0.0017;  
Matches 72; Conservative 343; Mismatches 310; Indels 10; Gaps 3;

QY 421 AGGACATCGAAGGATCATGATCCCTCTCTGTCGCTCAGGCGGTTAAGTCGATCGTCC 480  
DB 48 RKCSGCGKNTTTRKSKWYSASASGRTGSKWSSGYSKGMKKYKRSKRWGRGRRG 107  
QY 481 CAATTTTCTAGAAAAAGGTATGGAATTTGTCACAAGATGATGAGGATCGGCTGAGA 540  
DB 108 MERSRMWGRYRCARSGRAGSGRMMGKGRMSYMMWCMYARCGSCCKRKS--KGGS 165  
QY 541 AGGATATGCGCTGGGAGAGTCCGCGGTGAAAAGAGCAACACGACTCGAGACCGAAG 600  
DB 166 WGTCTRRGARGSGWSSGAKYKSGMSKRWMMSSCGRSGGRRSAYRYIGTSRKYTYK 225  
QY 601 GAGTCGATGTAAGGATGTTGGTTCGTCGACTACTCTGACGCTCATGCTCTTCGAGGAT 660  
DB 226 KWTYSASRCNRYATMTTSYKACSSYTWCRSKRSMMWCMYARCGSCCKRKS--KGGS 285  
QY 661 TGACTATAAGGAGCATCGCTCCAGAACCAAGCAATGAGCTCTATGCTGCTTTGTCG 720  
DB 286 MCTAYKKSYSYRWYRGYRGYRGYRGYRGYRGYRGYRGYRGYRGYRGYRGYRGYRGY 345  
QY 721 GACTTACCGATGGGTTGCTCTACCTACCTGAGCTCGTTCAAGGCTATCATGTGGGATTTG 780  
DB 346 RMMRSKCRWSKACYRWWRWMTERRRWAJKSSRTSRKKKCMWRKCKYKMRGYSR 405  
QY 781 TACCTTACTTCGAACTATGAACGAGAGACATGAGA-----TACCTTTGACTCAAGGAT 834  
DB 406 MRSCKRARMKRCRSGRWKNMGRGCMTCRMKSYGMMRWKSWRMSKRYMMRMYRWK 465  
QY 835 TAGCAGTTTCCGACGAGTGGGATCGAGCTTATGAGCAAAAGAGACGAGCGCTGCTG 894  
DB 466 KCSSTTWTGTTGMMGTGRCYKRSKMKRCRRRRWGRMYRWKRYNYSATYWR 525  
QY 895 GTCAGCTTCCGATCAGCTGTTGATAAAAAGGATGTTCAAGGTCGGGATATCTAAGTC 954  
DB 526 YCARKKYSYSAARKARCWYRGKYGYWAGMMWKRYMYKMMWYKRYKSKSWYCKMS 585  
QY 955 TCCTAGTCAGAGCAACATCGCC--GCCAACCTCGCTGATCTCAAAAGCTGTCGATGA 1012  
DB 586 YYASCMKSKARKAGMKCRKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMS 645  
QY 1013 GGAGGTACTCGCTCAGATCAGTAACTGTTATTGCTGGATATGAAACTCTTCCTCGACGT 1072  
DB 646 SBNKSYTCYWKMGSMKSTCTTMMYMSKYTYAKYGSYWRVRYRAWCMWMMWYNYRSY 705  
QY 1073 CTTGACATGGATGTTTCCAGACTCTCAGAGCAAAAGCGTTCAGGATTAACCTCGAGA 1132  
DB 1132 CTTGACATGGATGTTTCCAGACTCTCAGAGCAAAAGCGTTCAGGATTAACCTCGAGA 1132





Db 522 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 581

QY 1231 CTCCTAGTCGCTATACCGTGAATGCTTAAGGATGAAGACTTCA 1278

Db 582 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAAGATGTTGAGATCA 629

RESULT 6

ABK97694

ID ABK97694 standard; DNA; 1062 BP.

XX

AC ABK97694;

XX

DT 07-OCT-2002 (first entry)

XX

DE DNA encoding cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 7-13.

XX

KW Cytochrome P450; CYP3A43; CYP3A43; CYP3A4; CYP3A5; CYP3A7;

XX drug metabolism; drug design; drug screening; gene; ds.

XX

OS Homo sapiens.

XX

XX WO200244213-A1.

PN

XX

PD 06-JUN-2002.

XX

XX 28-NOV-2001; 2001WO-SE002631.

XX

XX 28-NOV-2000; 2000SE-00004366.

PR

XX 11-JUN-2001; 2001SE-00002061.

PR

XX (ZAPH/) ZAPHIROPOULOS P G.

PA

PA (FINT/) FINTA C.

XX

XX Zaphiropoulos PG, Finta C;

PI

XX

XX WPI; 2002-557532/59.

DR

DR P-PSDB; ABG68755.

XX

XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets

PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug

PT metabolism, in drug design and drug screening.

XX

XX Disclosure; Fig 13; 131pp; English.

PS

XX

CC The invention describes a cytochrome P450 protein (I) in which CYP3A43

CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub

CC fragments, variants and multiples of (I) having essentially the same

CC characteristics. (I) is useful as a medicament, and for evaluating drug

CC metabolism, in drug design, and drug screening, and in tests for

CC adjusting the dose of drugs. This sequence encodes a novel cytochrome

CC P450 protein

XX

SQ Sequence 1062 BP; 318 A; 239 C; 216 G; 289 T; 0 U; 0 Other;

Query Match 2.5%; Score 48.4; DB 6; Length 1062;

Best Local Similarity 50.7%; Pred. No. 0.011;

Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTCCTGGATG 1056

Db 413 AAGCTCTGCTCCGATCGAGCTCGTGGCCCAATCAATATCTTTATTTTCTGGCTATG 472

QY 1057 AAATCTTCTCGACGCTTTCATGATGTTTTCACCGACTCTCAGAAGACAAAGCCGTTTC 1116

Db 473 AAACACGAGCAGTGTCTCTCTTCATATGATGTAAGTGGCCACTCACCTGATGTC 532

QY 1117 AGATTAATCTCGAAGAAATTTGTCAGATCGACCGGATGCTTACGCTAGACGAAC 1176

Db 533 AGCAGAACTCGAGGAGAAATGATGTCAGTTTACCCCAATGAAGCCACCCACCTATG 592

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGAAGGAGTCTCTTCGCTAGACC 1230

Db 593 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 652

QY 1231 CTCCTAGTCGCTATACCGTGAATGCTTAAGGATGAAGACTTCA 1278

Db 653 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAAGATGTTGAGATCA 700

RESULT 7

ABK97693

ID ABK97693 standard; DNA; 1365 BP.

XX

AC ABK97693;

XX

DT 07-OCT-2002 (first entry)

XX

DE DNA encoding cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 4-13.

XX

KW Cytochrome P450; CYP3A43; CYP3A43; CYP3A4; CYP3A5; CYP3A7;

XX drug metabolism; drug design; drug screening; gene; ds.

XX

OS Homo sapiens.

XX

XX WO200244213-A1.

PN

XX

PD 06-JUN-2002.

XX

XX 28-NOV-2001; 2001WO-SE002631.

XX

XX 28-NOV-2000; 2000SE-00004366.

PR

XX 11-JUN-2001; 2001SE-00002061.

PR

XX (ZAPH/) ZAPHIROPOULOS P G.

PA

PA (FINT/) FINTA C.

XX

XX Zaphiropoulos PG, Finta C;

PI

XX

XX WPI; 2002-557532/59.

DR

DR P-PSDB; ABG68754.

XX

XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets

PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug

PT metabolism, in drug design and drug screening.

XX

XX Claim 9; Fig 12; 131pp; English.

PS

XX

CC The invention describes a cytochrome P450 protein (I) in which CYP3A43

CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub

CC fragments, variants and multiples of (I) having essentially the same

CC characteristics. (I) is useful as a medicament, and for evaluating drug

CC metabolism, in drug design, and drug screening, and in tests for

CC adjusting the dose of drugs. This sequence encodes a novel cytochrome

CC P450 protein

XX

SQ Sequence 1365 BP; 406 A; 299 C; 294 G; 366 T; 0 U; 0 Other;

Query Match 2.5%; Score 48.4; DB 6; Length 1365;

Best Local Similarity 50.7%; Pred. No. 0.013;

Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTCCTGGATG 1056

Db 716 AAGCTCTGCTCCGATCGAGCTCGTGGCCCAATCAATATCTTTATTTTCTGGCTATG 775

QY 1057 AAATCTTCTCGACGCTTTCATGATGTTTTCACCGACTCTCAGAAGACAAAGCCGTTTC 1116

Db 776 AAACACGAGCAGTGTCTCTCTTCATATGATGTAAGTGGCCACTCACCTGATGTC 835

QY 1117 AGATTAATCTCGAAGAAATTTGTCAGATCGACCGGATGCTTACGCTAGACGAAC 1176

Db 836 AGCAGAACTCGAGGAGAAATGATGTCAGTTTACCCCAATGAAGCCACCCACCTATG 895

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGAAGGAGTCTCTTCGCTAGACC 1230

Db 896 ATACTGTCTCAGATGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 955  
 QY 1231 CTCCTAGTCGGTATCTTAACCGTGAATGCTTAAGGATGAAGACTTCA 1278  
 Db 956 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGAGATGTTGATCA 1003

RESULT 8  
 AAQ87717  
 ID AAQ87717 standard; cDNA; 1512 BP.  
 XX  
 AC AAQ87717;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 14-NOV-1995 (first entry)  
 XX  
 DE Human cytochrome P450 molecular species 3A4 cDNA coding region.  
 XX  
 KW Human cytochrome P450; amplification; PCR; primer; expression vector;  
 KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;  
 KW carcinogen; mutagen; liver metabolism; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP644267-A2.  
 XX  
 PD 22-MAR-1995.  
 XX  
 PF 20-JUL-1994; 94EP-00111298.  
 XX  
 PR 20-JUL-1993; 93JP-00201120.  
 PR 21-JUL-1993; 93JP-00180246.  
 PR 30-JUL-1993; 93JP-00208279.  
 XX  
 PA (SUMO) SUMITOMO CHEM CO LTD.  
 XX  
 PI Hayaishi K, Sakaki T, Yabueaki Y, Komai K, Kaneko H, Nakatsuka I;  
 XX  
 DR WPI; 1995-116991/16.  
 DR P-PSDB; AAR72363.  
 XX  
 PT Evaluation of safety of a chemical cpd. - using recombinant yeast  
 PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.  
 XX  
 PS Example; Page 31-33; 124pp; English.  
 XX

The nucleotide sequence of the cDNA coding region for the human cytochrome P450 species 3A4. The gene encodes a protein of 503 amino acids. The cDNA was amplified by PCR using the primers AAQ87743-6. The product was cloned into the yeast expression vectors pAAH5N or pAHRH to produce the vectors p3A4 for the expression of the cytochrome P450 alone or p3A4R for co-expression with the yeast NADPH-P450 reductase. The vectors are used in a method for evaluating the safety of a chemical compound by reacting the chemical compound with recombinantly produced human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2B1 (AAQ87716), or 3A4 or their auxiliary species and variants (AAQ87718-32), and yeast NADPH-P450 reductase, either as a fused protein or in cell extracts, and analysing the resulting metabolite to assess the safety of the chemical compound. The method is useful for determining whether the chemical compound, or its metabolite, will be converted into a carcinogenic or mutagenic form through metabolism in the liver. (Updated on 25-MAR-2003 to correct PN field.)

XX  
 SQ Sequence 1512 BP; 441 A; 330 C; 330 G; 411 T; 0 U; 0 Other;  
 Query Match 2.5%; Score 48.4; DB 2; Length 1512;  
 Best Local Similarity 50.7%; Pred. No. 0.013;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGGAGTACTCCGTCAGATCAGTAACCTGTTATTTCTGGATATG 1056  
 Db 863 AAGCTCTGTCGATGAGGAGTCTGCGTCCGATCATCTTTATTTTCTGGCATG 922

QY 1057 AAATCTTTCGACAGTCTTGACATGGATGTTTACCGACTCTCAGAAGACAAGCCGTTTC 1116  
 Db 923 AAACACGACGAGTGTCTCTCTCTCTCATTTATGATGACCTGCGCACTACCCCTGATGTC 982  
 QY 1117 AGGATAAACTTCGAGAAAGAAATTTGTTCAGATCGACACGGATATGCTTACGCTAGCAGAAC 1176  
 Db 983 AGCAGAACTTCGAGGAGGAAATTTGATGACGTTTACCCCAATAGGCACCACCACCTATG 1042  
 QY 1177 TTAATGGGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGTTCTAGACC 1230  
 Db 1043 ATACTGTCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1102  
 QY 1231 CTCCTAGTCCGTATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278  
 Db 1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGAGATGTTGATCA 1150

RESULT 9  
 AAT17399  
 ID AAT17399 standard; cDNA; 1512 BP.  
 XX  
 AC AAT17399;  
 XX  
 DT 01-AUG-1996 (first entry)  
 XX  
 DE Human derived cytochrome P4503A4 cDNA.  
 XX  
 KW Human derived cytochrome; P4503A4; commercial cDNA library; yeast;  
 KW transfection; recombinant production; expression vector; mammal;  
 KW immunisation; sensitisation; antibody; determination; detection;  
 KW non-cross reactive; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1512  
 FT /\*tag= a  
 XX  
 PN JP08027197-A.  
 XX  
 PD 30-JAN-1996.  
 XX  
 PF 13-JUL-1994; 94JP-00161552.  
 XX  
 PR 13-JUL-1994; 94JP-00161552.  
 XX  
 PA (SUMO) SUMITOMO CHEM CO LTD.  
 XX  
 WIPI; 1996-136338/14.  
 DR P-PSDB; AAR81464.  
 XX  
 PT Antibody recognising human derived cytochrome P4502A4 - allows specific  
 PT detection of cytochrome P450 species in humans.  
 XX  
 PS Example 1; Page 10-12; 13pp; Japanese.  
 XX  
 CC The present sequence encodes the human derived cytochrome (HDC) P4503A4,  
 CC which was obtained from a commercial cDNA library. Yeast were transfected  
 CC with an expression vector containing the HDC cDNA, cultured and then  
 CC disrupted to give a microsomal fraction. The HDC was purified from the  
 CC fraction, and used to immunise and sensitise a mammal. Blood was drawn  
 CC from the mammal, and an anti-HDC antibody isolated. The antibody obtained  
 CC recognises HDC P4503A4, partic. at a serum dilution rate of 1:10000, and  
 CC is substantially without cross reaction to other HDC P450 spp  
 XX  
 SQ Sequence 1512 BP; 441 A; 330 C; 330 G; 411 T; 0 U; 0 Other;  
 Query Match 2.5%; Score 48.4; DB 2; Length 1512;  
 Best Local Similarity 50.7%; Pred. No. 0.013;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGGAGTACTCCGTCAGATCAGTAACCTGTTATTTCTGGATATG 1056  
 Db 863 AAGCTCTGTCGATGAGGAGTCTGCGTCCGATCATCTTTATTTTCTGGCATG 922

Db 863 AAGCTCTCTCGATCTGGAGCTCGTGGCCCAATCAATATCTTTATTTTGGCTATG 922  
 QY 1057 AAACCTTCTTCGACAGCTTTGACATGGATGTTTCCAGGACAAAGCGTTTC 1116  
 Db 923 AAACCAAGCAGAGTGTCTCTCTCTTCAATATGATGATGAAGTGGCACTCCACCTGATGTC 982  
 QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTTACCTAGCTAGACGAC 1176  
 Db 983 AGCAGAACTCGAGAGGAAATTTGATGAGATTTTACCAATAGGACACACCCACCTATG 1042  
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230  
 Db 1043 ATACTGTGCTACAGATGAGTATCTTACATGCTGGTGAATGAAGACGCTCAGATTATTC 1102  
 QY 1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278  
 Db 1103 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAAAGATGTTGAGATCA 1150

RESULT 10  
 AAT28383  
 ID AAT28383 standard; DNA; 1512 BP.  
 XX  
 AC AAT28383;  
 XX  
 DT 11-OCT-1996 (first entry)  
 XX  
 DE Human cytochrome P450 molecular species 3A4 gene.  
 XX  
 KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
 KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
 KW evaluation; safety; fusion protein; metabolite; detoxification;  
 KW carcinogenic; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JF08056695-A.  
 XX  
 PD 05-MAR-1996.  
 XX  
 PF 15-JUL-1994; 94JP-00164184.  
 XX  
 PR 20-JUL-1993; 93JP-00201120.  
 PR 30-JUL-1993; 93JP-00208279.  
 PR 17-JUN-1994; 94JP-00136053.  
 XX  
 (SUMO ) SUMITOMO CHEM CO LTD.  
 XX  
 WPI: 1996-182311/19.  
 DR P-PSDB; AAR93170.  
 XX

Novel method for the evaluation of the safety of a cpd. - using a human cytochrome P450 and yeast NADPH reductase to determine whether the analyte cpd. is detoxified or metabolised to a carcinogen.

Example 1; Page 24-26; 74pp; Japanese.

This is the nucleotide sequence of the human cytochrome P450 molecular species 3A4 gene which encodes a protein of 503 amino acids. The gene was amplified from a human liver derived cDNA library as 2 fragments of 0.6 and 0.9 kb using primers AAT28383-6. The prod. was cloned into the yeast expression vector pAAH5N to generate plasmid p3A4 for prodn. of the cytochrome only or into the vector pAHR to generate the plasmid p3A4R for co-prodn. with the yeast NADPH-P450 reductase. The sequence is placed under control of the yeast ADH gene promoter and terminator. The vectors are used in a method for evaluating the safety of a cpd. by reacting the test cpd. with recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380), 2C9 (AAT28381), 2E1 (AAT28382), 3A4 or their variants (AAT28384-98) together with yeast NADPH-P450 reductase (either as a fused protein or as a cell extract) and analysing the resultant metabolite. The cpd. is considered "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it is not detoxified or is metabolised to a carcinogenic cpd

XX Sequence 1512 BP; 441 A; 330 C; 330 G; 411 T; 0 U; 0 Other;  
 SQ  
 Query Match 2.5%; Score 48.4; DB 2; Length 1512;  
 Best Local Similarity 50.7%; Pred. No. 0.013;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
 QY 997 AAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGATATG 1056  
 Db 863 AAGCTCTCTCGATCTGGAGCTCGTGGCCCAATCAATATCTTTATTTTGGCTATG 922  
 QY 1057 AAACCTTCTTCGACAGCTTTGACATGGATGTTTCCAGGACAAAGCGTTTC 1116  
 Db 923 AAACCAAGCAGAGTGTCTCTCTCTTCAATATGATGATGAAGTGGCACTCCACCTGATGTC 982  
 QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTTACCTAGCTAGACGAC 1176  
 Db 983 AGCAGAACTCGAGAGGAAATTTGATGAGATTTTACCAATAGGACACACCCACCTATG 1042  
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230  
 Db 1043 ATACTGTGCTACAGATGAGTATCTTACATGCTGGTGAATGAAGACGCTCAGATTATTC 1102  
 QY 1231 CTCCTAGTCCGATGCTTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278  
 Db 1103 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAAAGATGTTGAGATCA 1150

RESULT 11  
 ABK97692  
 ID ABK97692 standard; DNA; 1512 BP.  
 XX  
 AC ABK97692;  
 XX  
 DT 07-OCT-2002 (first entry)  
 XX  
 DE DNA encoding cytochrome P450 CYP3A4 exon 1-CYP3A4 exon 2-13.  
 XX

KW Cytochrome P450; CYP3AP1; CYP3AP2; CYP3A43; CYP3A4; CYP3A5; CYP3A7;  
 KW drug metabolism; drug design; drug screening; gene; ds.  
 XX Homo sapiens.  
 OS  
 PN WO200244213-A1.  
 XX  
 PD 06-JUN-2002.  
 XX  
 PF 28-NOV-2001; 2001WO-SE002631.  
 XX  
 PR 28-NOV-2000; 2000SE-00004366.  
 PR 11-JUN-2001; 2001SE-00002061.  
 XX  
 (ZAPH/) ZAPHIROPOULOS P G.  
 PA (FINT/) FINTA C.  
 XX  
 PI Zaphiropoulos PG, Finta C;  
 XX  
 WPI: 2002-557532/59.  
 DR P-PSDB; ASG68753.  
 XX

Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug metabolism, in drug design and drug screening.  
 Claim 9; Fig 11; 131pp; English.  
 CC The invention describes a cytochrome P450 protein (I) in which CYP3A43 exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub fragments, variants and multiples of (I) having essentially the same characteristics. (I) is useful as a medicament, and for evaluating drug metabolism, in drug design, and drug screening, and in tests for adjusting the dose of drugs. This sequence encodes a novel cytochrome P450 protein

XX SQ Sequence 1512 BP; 447 A; 324 C; 328 G; 413 T; 0 U; 0 Other;  
Query Match 2.5%; Score 48.4; DB 6; Length 1512;  
Best Local Similarity 50.7%; Pred. No. 0.013;  
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
QY 997 AAAAGCTGTCGATGAGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056  
DB 863 AAGCTCTGTCGATGTCGAGCTCGTGGCCCAATCAATTATCTTTATTGCTGGCTATG 922  
QY 1057 AAACTTCTTCGACAGCTTTGACATGGATGTTTCCAGCACTCTCAGAGACAAAGCCGTTTC 1116  
DB 923 AAACACGAGCAGTGTCTCTCCITCATATGATGTAAGTGGCCACTCACCTGATGTC 982  
QY 1117 AGGATAAACTTCGAGAGAAATTTGTCAGATCGACACGGATATGCCCTACGCTAGACGAAC 1176  
DB 983 AGCAGAAACTGCGAGGAGAAATTTGATGAGTGTATACCCCAATAGGCAACCCACCTATG 1042  
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230  
DB 1043 ATACTGTCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1102  
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATCTTAAAGGATGAAGACTTCA 1278  
DB 1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1150  
RESULT 12  
ACF06054  
ID ACF06054 standard; cDNA; 1512 BP.  
XX  
AC ACF06054;  
XX  
DT 22-SEP-2003 (first entry)  
DE Human cytochrome P450 3A4 encoding cDNA.  
XX  
KW Human; protein array; protein moiety; phenotype; drug discovery;  
KW naturally occurring variant; pharmacogenomic; diagnostic;  
KW parallel analysis; tumour suppressor; p53; cytochrome P450; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2003048768-A2.  
XX  
PD 12-JUN-2003.  
XX  
PF 05-DEC-2002; 2002WO-CB005499.  
XX  
XX 05-DEC-2001; 2001US-0335806P.  
PR 16-SEP-2002; 2002US-0410815P.  
XX  
XX (SENS-) SENSE PROTEOMIC LTD.  
XX  
PI Boutell JM, Godber BLJ, Hart DJ, Blackburn JD;  
XX  
XX WPI; 2003-569063/53.  
DR P-PSDB; ABR82024.  
XX  
PT New protein array, useful for determining the phenotype of a naturally  
PT occurring variant of a DNA sequence of interest, comprises a surface upon  
PT which at least two protein moieties are deposited.  
XX  
XX Example 5; Fig 11A; 84pp; English.  
XX  
XX The present invention describes a protein array comprising a surface upon  
XX which at least two protein moieties are deposited at spatially defined  
XX locations, where the protein moieties are naturally occurring variants of  
XX a DNA sequence of interest. Also described: (1) making a protein array;  
XX (2) screening a set of protein moieties for molecules that interact with  
XX one or more proteins; and (3) simultaneously determining the relative  
XX properties of members of a set of protein moieties. The protein array can

CC be used for determining the phenotype of a naturally occurring variant of  
CC a DNA sequence of interest. The protein array is useful for drug  
CC discovery, pharmacogenomics and diagnostics. The protein array allows the  
CC parallel analysis of closely related proteins with a sensitivity that is  
CC at least comparable to existing methods, if not better, with small  
CC volumes of potentially expensive ligands, and in a quantitative,  
CC comparative functional analysis manner not previously possible. ACF06000  
CC to ACF06056 and ABR81975 to ABR82026 represent sequences used in the  
CC exemplification of the present invention  
XX SQ Sequence 1512 BP; 441 A; 330 C; 330 G; 411 T; 0 U; 0 Other;  
Query Match 2.5%; Score 48.4; DB 8; Length 1512;  
Best Local Similarity 50.7%; Pred. No. 0.013;  
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
QY 997 AAAAGCTGTCGATGAGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056  
DB 863 AAGCTCTGTCGATGTCGAGCTCGTGGCCCAATCAATTATCTTTATTGCTGGCTATG 922  
QY 1057 AAACTTCTTCGACAGCTTTGACATGGATGTTTCCAGCACTCTCAGAGACAAAGCCGTTTC 1116  
DB 923 AAACACGAGCAGTGTCTCTCCITCATATGATGTAAGTGGCCACTCACCTGATGTC 982  
QY 1117 AGGATAAACTTCGAGAGAAATTTGTCAGATCGACACGGATATGCCCTAGTAGACGAAC 1176  
DB 983 AGCAGAAACTGCGAGGAGAAATTTGATGAGTGTATACCCCAATAGGCAACCCACCTATG 1042  
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230  
DB 1043 ATACTGTCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1102  
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATCTTAAAGGATGAAGACTTCA 1278  
DB 1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1150  
RESULT 13  
ABN95088  
ID ABN95088 standard; DNA; 2011 BP.  
XX  
AC ABN95088;  
XX  
DT 13-AUG-2002 (first entry)  
DE Gene #1586 used to diagnose liver cancer.  
XX  
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KW metastatic liver tumour; cytostatic; expression profile; disease state;  
KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX  
OS Homo sapiens.  
XX  
XX WO200229103-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 02-OCT-2001; 2001WO-US030589.  
XX  
XX 02-OCT-2000; 2000US-0237054P.  
XX  
XX (GENE-) GENE LOGIC INC.  
XX  
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX WPI; 2002-426119/45.  
XX  
XX Diagnosing and detecting the progression of liver cancer, hepatocellular  
XX carcinoma or metastatic liver tumor in a patient, involves detecting the  
XX level of expression of two or more genes in a liver tissue sample.  
XX  
XX Claim 1; SEQ ID NO 1586; 298pp; English.  
XX



PA (GENE-) GENE STREAM PTY LTD.  
 XX Dally JM;  
 PI  
 XX  
 XX WPI; 2003-093021/08.  
 DR P-PSDB; ABUS7260.  
 DR  
 XX  
 XX New transgenic non-human animal expressing a foreign polypeptide  
 PT associated with drug behavior and/or metabolism, useful for studying the  
 PT behavior and/or metabolism of a drug in other animals.  
 XX  
 XX  
 PS Example 2A; Page 296-299; 408pp; English.  
 XX  
 XX This invention relates to a transgenic non-human animal which may be used  
 CC for assessing the behaviour and/or metabolism of a drug in another animal  
 CC and which expresses a foreign polypeptide associated with drug behaviour  
 CC and/or metabolism. The invention also comprises a nucleic acid construct  
 CC for use in producing the above transgenic non-human animal and a method  
 CC of assessing the metabolism and/or behavior of a drug in an animal of  
 CC interest, comprising administering a test agent to the transgenic animal  
 CC and conducting analytical tests to determine drug metabolism and/or  
 CC behaviour. The transgenic animal is useful in studying drug metabolism  
 CC and/or behaviour in other animals. The nucleic acid construct is useful  
 CC in producing the above transgenic animal and the methods are used for  
 CC producing, breeding and using transgenic animals for pharmacological  
 CC (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological  
 CC studies. Nucleic acid sequences used within the invention are serum  
 CC albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine  
 CC diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins  
 CC and (MRP's). The present sequence represents a DNA sequence used to  
 CC create a transgenic animal within the scope of the invention  
 XX  
 SQ Sequence 2764 BP; 823 A; 582 C; 588 G; 771 T; 0 U; 0 Other;  
 Query Match 2.5%; Score 48.4; DB 7; Length 2764;  
 Best Local Similarity 50.7%; Pred No. 0.018;  
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;  
 QY 997 AAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTCTTATTTGCTGGATATG 1056  
 Db |||||  
 QY 967 AAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTCTTATTTGCTGGATATG 1026  
 Db |||||  
 QY 1057 AACTTCTTCGACAGTCTTGATCGATGTTTACCGACTCTCAGAGACAAAGCCGTTTC 1116  
 Db |||||  
 QY 1027 AACCCAGAGCAGTGTCTCTCTTCATTATGATGAACTGGCCACTCACCTGATGTCC 1086  
 Db |||||  
 QY 1117 AGGATAAATTCGAGAAGAAATTTCTCAGATCGACACGGATATGCTACCGTAGACGAAC 1176  
 Db |||||  
 QY 1087 AGCAGAACTCGAGGAGGAAATTTGATGCGATTTTACCCCAATAAGGCACACCCACTATG 1146  
 Db |||||  
 QY 1177 TTAATCGGTTC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGTAGACC 1230  
 Db |||||  
 QY 1147 ATACTGTGTACAGATGGAGTATCTTGTGATGCTGGTGAATGAACCGCTCAGATTATTC 1206  
 Db |||||  
 QY 1231 CTCTAGTCCGTATGTTAAACCGTGAATGTTAAAGGATGAAGACTTCA 1278  
 Db |||||  
 QY 1207 CAATGCTATGAGACTTGAGAGGGTCTGCAGAAAAGTGTGAGATCA 1254  
 Db |||||

Search completed: April 6, 2004, 14:11:28  
 Job time : 841 secs